HELSINKI UNIVERSITY OF TECHNOLOGY
Department of Electrical and
Communications Engineering

Tiina Lindh-Knuutila

Simulating the Emergence of a Shared Conceptual System in a Multi-Agent Environment

Master's Thesis submitted in partial fulfillment of the requirements
for the degree of Master of Science in Technology

Espoo, October 10, 2005

Supervisor: Docent Timo Honkela
Instructor: Krista Lagus, Dr.Sc.(Tech.)

TEKNILLINEN KORKEAKOULU DIPLOMITYÖN TIIVISTELMÄ

Tekijä:	Tiina Lindh-Knuutila		
Työn nimi:	Yhteisen käsitejärjestelmän muodostumisen simulointi moniagenttiympäristössä		
Päivämäärä:	10.10.2005	**Sivuja:**	62+7
Osasto:	Sähkö- ja tietoliikennetekniikka		
Professuuri:	T-61 Informaatiotekniikka		
Työn valvoja:	Dosentti Timo Honkela		
Työn ohjaajat:	TkT Krista Lagus		

Ihmiset kykenevät välittämään merkityksiä toisilleen käyttämällä yhteisiä, sovittuja symboleja eli sanoja. Tässä diplomityössä tutkitaan käsitteiden ja sanojen välisten yhteyksien muodostumista, eli sitä miten kielen käyttäjä eli agentti oppii uusien sanojen merkityksen, ja sitä miten yhteinen kieli syntyy agenttipopulaatiossa.

Diplomityössä agentin käsitekarttaa mallinnetaan itseorganisoituvan kartan avulla. Käsitteinä pidetään tässä ohjaamattoman oppimisen avulla itseorganisoituvalle kartalle syntyviä alueita. Kartan voidaan ajatella vastaavan käsiteavaruuden yhtä tasoa. Kielen oppimista mallinnetaan kielipelien, erityisesti havaintoihin perustuvan kielipelin, avulla simuloidussa agenttipopulaatiossa. Työssä toteutetaan agenttisimulaatioympäristö, jota testataan erilaisia parametriarvoja käyttäen. Kokeiden tulokset vahvistavat, että simulaation edetessä agentit oppivat kommunikoimaan onnistuneesti yhteistä, emergoituvaa sanastoa käyttäen.

Avainsanat: kielipelit, kielen oppiminen, käsitemallinnus, agenttisimulaatio, käsiteavaruus.

HELSINKI UNIVERSITY OF TECHNOLOGY		ABSTRACT OF THE MASTER'S THESIS
Author:	Tiina Lindh-Knuutila	
Name of the thesis:	Simulating the Emergence of a Shared Conceptual System in a Multi-Agent Environment	
Date:	October 10, 2005	**Number of pages:** 62+7
Department:	Electrical and Communications Engineering	
Professorship:	T-61 Computer and Information Science	
Supervisor:	Docent Timo Honkela	
Instructor:	Krista Lagus, Dr.Sc.(Tech.)	

In this Master's Thesis work, the emergence of associations between concepts and words is studied. The important questions are how a language learner, or an agent, learns the meaning of new words, and how an agreement on the use of words is reached in a community of agents.

The Self-Organizing Map is used as a model of an agent's conceptual map, and concepts are thought to be areas formed in a self-organizing map based on unsupervised learning. The map may be thought to be an equivalent of a domain in a Conceptual Space. The language acquisition process is modeled in a population of simulated agents by using a series of language games, called the observational games. For the experiments, an agent simulation framework is implemented and tested with different parameters. The results of the experiments verify that the agents learn to communicate successfully and a shared lexicon emerges.

Keywords: language games, language acquisition, conceptual modeling, agent simulation, conceptual spaces

Acknowledgements

This Master's Thesis has been done in the Laboratory of Computer and Information Science at Helsinki University of Technology.

First of all, I want to thank my supervisor, Docent Timo Honkela, for giving me the opportunity to do my Master's Thesis on this interesting topic as well as for the ideas and supportive discussions throughout the process. My gratitude also goes to the instructor of this work, Dr. Krista Lagus, for encouragement and valuable ideas and comments throughout the work.

I would also like to thank the past and present members of the Computational Cognitive Systems research group for helpful discussions. Especially I would like to thank Jaakko and Matti for always having time to answer my questions and for proof-reading the manuscript.

Finally, I would like to thank my parents for love and support and Marko for love, encouragement and late-night discussions.

Espoo, October 10, 2005

Tiina Lindh-Knuutila

Contents

Abbreviations

AI	Artificial Intelligence
BMU	Best Matching Unit
CS	Communication Success
CST	Conceptual Spaces Theory
GG	Guessing Game
ILM	Iterated Learning Model
LOT	Language Of Thought
MN	Mirror Neuron
OG	Observational Game
SG	Selfish Game
SOM	Self-Organizing Map

List of Figures

Chapter 1

Introduction

Humans are able to convey meanings to each other by using common symbols, words. Language is often considered to be a hallmark of intelligence [27]. In this Master's Thesis, the emergence of associations between concepts and words is studied with help of a computer simulation. The important question that is addressed here is how a language learner, an agent, acquires the meaning of new words.

Acquisition of concepts and language goes hand in hand. Because of this, the Thesis includes discussion on concepts and their formation. What are the views on concepts? How could their acquisition be modeled? A discussion on the theme is very important. The concepts and the modeling of concepts are discussed in Chapter 2.

Language acquisition and the emergence of common vocabulary are in the scope of this thesis. Chapter 3 concentrates on these themes. The nativist versus non-nativist debate on language origins is briefly touched, but mainly we are interested in language acquisition and its computational modeling.

In this work, the acquisition of language is modeled with help of multi-agent simulations using the language game approach. The approach used in the Thesis is presented in more detail in Chapter 4. It follows the research on observational game by Vogt [49]. In this Thesis, the concepts are modeled as areas formed in a Self-Organizing Map [18] due to unsupervised learning. The map may be seen as representing a domain in a conceptual space, introduced by Gärdenfors [7]. In this way, the Self-Organizing Map is used as a model for an agent's ontology.

The experimental settings and the results for the experiments are presented in Chapter 5. The purpose of the experiments is to confirm the hypothesis that after enough language games the agents are able to communicate with an emerging shared language, when the Self-Organizing map is used as a model for conceptual spaces.

In this thesis, computer simulations of multiple agents are used. It seems thus appropriate to address the question on the role of computer simulations in cognitive systems research. In general, it can be seen that the autonomous agents research in cognitive sciences has two goals [52]: The engineering point of view is concerned mostly with the design and construction of artefacts. The scientific point of view tries to explain how natural systems work. In this work an agent simulation environment is constructed based on observations on language acquisition process in humans. Some attempts are made to capture real aspects of language in these simulations. But even when a language emerges, the simulation environment is only a simple piece of (wo)man-made program dependent on the initial settings. Thus, the purpose is only to show that the process might be "somewhat like this".

The use of a simulated environment raises another question. In [36], strong scepticism toward the use of agent simulations is presented:

> Computational models and artificial models [...] must be clearly distinguished. For example, it is possible to build a computational model of how a bird flies, which amounts to a simulation of the environment around the bird, a simulation of the aerodynamics of the body and the wings, a simulation of the pressure differences caused by movement of

> the wings, etc. Such a model is highly valuable but would however not be able to fly. [...] Very often results from simulation only partially carry over to artificial systems. When constructing a simulation, one selects certain aspects of the real world that are carried into the virtual world. But this selection may ignore or overlook essential characteristics which play a role unknown to the researcher.

When constructing artificial systems, should not we then rather build real systems equipped with means to explore a real (physical) environment? One cannot deny the importance of 'real' robotic experiments in this field. But as Ziemke [52] argues, simulations have an important, complementary role in the field of research: In many cases they allow for more extensive, systematic experimentation.

A thorough overview on the advantages and disadvantages on robotic experiments and simulated environments is presented in [27]. The main points are that simulations are in principle fast, cheap and flexible. Additionally, the social dimensions of (multiple agents) are easier to create than in robotics and the experimenter does not need to be present all the time, and the debugging is easier. But as Steels has pointed out in [36], reliable and good simulations of physical systems are hard to create.

The structure of the remaining thesis is the following. In Chapter 2 theoretical background of the concepts and their acquisition is presented. Chapter 3 discusses language, language acquisition and naming game or language game models by Steels [37] and others. Chapter 4 presents the approach taken in this Master's Thesis. Chapter 5 contains the description of the experiments and the results. Chapter 6 concludes and an outline for the future work around the theme is given.

Chapter 2

Conceptual modeling

What is a concept? How should concepts be modeled? Are they innate, should they be learned? This chapter tries to find answers to these questions by presenting some views on concepts.

There is a constant debate on the principles behind the conceptual systems. The supporters of the Language of Thought hypothesis (LOT), Jerry Fodor being one of the most famous of them, argue that to have a conceptual representation is to possess some kind of expression which is made of concatenated symbols [1]. A cognitive agent is seen as some kind of logic machine that operates using these symbols (mind-as-a-machine) [20]. These symbols then form a mental language or mentalese, as it is sometimes called, which is seen as innate. Concept learning is then re-testing of hypotheses available already at birth. For others, this kind of view on the innateness of concepts is already slightly disturbing. For many people (including the author), it seems quite counter-intuitive that, e.g., the concept of carburetor would be innate (example from [21]).

The symbolic approach has generally been adopted in traditional AI simulations. In those systems, the meanings used by the agents were simply other, hand coded symbols. The intelligent behavior was then manipulating these symbols in a rule-based manner. Here, symbols mean arbitrary tokens that can be manipulated by rules.

Now we get to the problematic issue: What are these mental symbols then? From where does any meaning come to them? This question is called *the symbol grounding problem* and it was compressed to the following form by Harnad [10, p. 335]:

> How can the semantic interpretation of a formal symbol system be made intrinsic to the system, rather than just parasitic on the meanings in our heads? How can the meanings of the meaningless symbol tokens, manipulated solely on the basis of their (arbitrary) shapes, be grounded in anything but other meaningless symbols?

Harnad illustrated this problem by constructing his version of Searle's [33] symbol grounding problem. In this problem, a person is trying to learn Chinese from a Chinese-Chinese dictionary only, thus replacing symbols with other symbols found from the dictionary ad infinitum.

Harnad's solution to the symbol grounding problem was to ground symbolic representations bottom-up from two kinds of non-symbolic representations: 1) iconic representations that are analogs of proximal sensory projections and 2) categorical representations that are learned by innate feature detectors picking up the invariant features of object and event categories from the sensory projections. The elementary symbols would then be names for these categories and higher-order symbolic representations would be grounded in the elementary symbols.

A strong opponent to the LOT view is Lakoff who is working with the embodied cognitive models [20]. The principle of embodiment is that the meaning is grounded in bodily experiments. In [20, p. 206], he offers strong criticism on symbolical concept systems:

> Cognitive models that are embodied are not made up merely of items in

> an artificial language. [...] In objectivist accounts such [bodily] experiences are simply absent. It is as though human beings did not exist, and their language and its (not their) meanings existed without any beings at all.

Lakoff's view is that embodied cognitive models structure thought and they are used in forming categories and in reasoning. The cognitive models characterize the concepts, which are used via the embodiment of the models. A further note on the embodiment is that most cognitive models are embodied with respect to use. Abstract conceptual structures are indirectly meaningful: They are understood because of their systematic relationship to directly meaningful structures.

Of course, the embodiment approach as such does not provide help in finding out how one gets from the continuous sensory signals to the symbolic level of words. Anyhow, it provides a link for the meaning creation and acquisition.

2.1 Different views on concepts

In this section, three views on concepts are presented in more detail. First of them is the Classical View. The second, the Prototype Theory, began from Eleanor Rosch's observations on prototype effects. The third is the Conceptual Spaces Theory developed by Peter Gärdenfors. For a thorough review on different views on concepts including some that are not presented here, consider Laurence and Margolis' review [21]. In the following, some basic definitions related to conceptual systems are presented.

In general, concepts are seen as mental particulars in the field of linguistics and cognitive science. This is not agreed by all. For example, it is usual to think concepts as abstract entities in the field of philosophy [21]. Very often theories of concepts concentrate on the notion of lexical concepts — i.e., concepts that correspond to lexical items in natural language. In this work, the word concept is taken as a means for specifying a relationship between world and language.

A category is a group of objects that are considered equivalent, and generally a category has a name, e.g., 'dog' or 'animal'. [31]. Rosch presents two principles according to which the category systems function. These principles are 1) cognitive economy and 2) perceived world structure. The function of the category systems is to provide maximum information regarding the world with the least cognitive effort.

For an individual, it is useful in many ways to be able to distinguish perceived items from other items. This sorting to categories is often called *categorization*. It has a major role in perception, thinking and language, and it is probably significant in motor performance as well [9].

By categorical perception, it is commonly meant that the differences among items belonging to the same category are diminished and the differences between items falling to different categories are magnified. This phenomenon has been demonstrated, e.g., for speech [9].

The purpose of the conceptual system is to *interpret* the world, not merely to record it as a video recorder does [2]. Thus, a human categorizes components from for example a photo, which is beyond the capabilities of a recorder. Furthermore, a human is capable of drawing inferences from these categorized components.

2.1.1 Classical Theory

In the Classical Theory of concepts, the structure of concepts is seen as a set of definitions [21]. According to this view, for example, a concept for 'bird' could include the following set of definitions: 'has wings', 'can fly', 'lays eggs', etc. that are necessary and sufficient conditions for something to be a bird. This kind of notion on concepts has a long history in philosophy (e.g., [22]).

This kind of view to concepts is quite problematic. Many problems faced by the Classical Theory can be listed [21]: For instance, there are only few examples of (well) defined concepts, and even lexical concepts do not show effects of definitional structure in psychological experiments. Wittgenstein's writings [51] on family resemblance were among the first to question the classical view. He pointed out that there are categories like 'games' that do not fit into this kind of theory. There are all kinds of games but the games do not share properties that are common to all of them. What makes it possible for us to have a category for games is family resemblance: They are similar to each other in a variety of ways instead of them all having certain common definitions.

It is also possible to have a concept in spite of massive ignorance or error. We are able to 'have a concept' even if we are mistaken about the properties we think its instances to have. Laurence and Margolis [21] take diseases as an example. People used to believe the diseases were caused by evil spirits, or, in case of a physical explanation, 'bad blood'. Nowadays it is however believed that these people were wrong about the nature of such diseases, or that they made a coarse-grained categorization into physiological and psychological domain. But saying this supposes that we are still talking about the same concept. Thus, their most fundamental beliefs or definitions could not be correct, and the matter of possessing a concept cannot be knowing the necessary and sufficient conditions for its application.

Another problem is that some concepts are fuzzy, i.e., boundaries of the categories are not sharp. According to the classical theory, in which the categorization should always be determinate, fuzzy categorization should not be possible.

A further problem are typicality effects, which proved to be the most influential argument against the Classical Theory. In experiments, people judge some instances of a category to be better examples of a concept than others. For example, a sparrow is judged as a more typical example of the category 'bird' than a penguin [21].

A newer version of the classical theory is the neoclassical theory, in which the definitions encoding the concept are thought to be partial [21].

2.1.2 Prototype Theory

The Prototype Theory emerged in the 1970s as an alternative to the classical theory to include the experimental findings of typicality effects of concepts [21]. The core idea of the prototype theory is that most lexical concepts are complex representations, whose structure encodes a statistical analysis of the properties their members tend to have.

Prototype theory has its own problems as well. Some researchers argue that the existence of prototypes tells nothing about concepts, since well defined concepts also exhibit typicality effects. The problem of ignorance and error is as much a problem in prototype theory as it was a problem in the classical theory.

Many concepts do not seem to have prototypes at all. It seems that there are some concepts for which people fail to represent any central tendencies at all. Laurence and Margolis [21] give some examples of them. There are concepts that are not instantiated at all, e.g., 31st century invention, others that have too heterogeneous extensions, e.g., objects that weigh more than a gram, and some others that seem to be too abstract, e.g., belief. As a related problem they state that it is perfectly possible to have a concept without knowing a prototype for it, even if others who possess the concept do.

Rosch herself argues [31] that empirical findings of prototypicality effects have been confused with theories of processing: They seem only to constrain, but not specify, representation and process models. She also writes that prototypes appear to be just those members of a category that most reflect the redundancy structure of the category as a whole. Following this line of thought, it could be said that the aforementioned concepts are similar in such a way that there are not enough redundancy structure that could be reflected in a form of prototypes.

There is also the problem of how compositional concepts are formed. In many cases a prototype for a complex concept is not a prototype for the constituents of that complex concept [21].

2.1.3 Conceptual Spaces

The Conceptual Spaces Theory (CST) [7] was developed by Gärdenfors in order to be able to model conceptual representations in a cognitive framework. According to the theory, concepts could be modeled as geometrical areas in a multidimensional conceptual space rather than as symbols or connections among neurons. Or rather, there are three different levels of representation, in which each of these approaches is suitable: symbolic, conceptual and sub-conceptual.

A conceptual space is built upon geometrical structures based on a number of quality dimensions. Concepts are not independent of each other but can be structured into domains, e.g., concepts for colors in one domain, spatial concepts in other domain. These quality dimensions represent various 'qualities' of objects. Gärdenfors also claims that the quality dimensions of conceptual spaces are independent of and more fundamental than the symbolic representations.

A conceptual space consists of a class $D_1, D_2, \ldots, D_n$ of quality dimensions. A point in the space is represented by a vector $v = [d_1, d2, \ldots, d_n]$ [7]. Temperature, weight, brightness, and the spatial dimensions height, width and depth are listed as possible quality dimensions perceivable with the human sensory system. Gärdenfors points out that the metrics vary according to the perceiver. For example, the temperature is not necessarily perceived similarly by everybody and the height is perceived differently from different distances. It is also pointed out that there is not, in general, a unique way of choosing a dimension to represent a particular quality, but various possibilities. As Gärdenfors points out, even though the representations of the world and our perceptions of it vary, scientific representations could be used in construction of an artificial system. When constructing an artificial system, the input on different sensors is described in terms of scientifically modeled dimensions.

He sees the basic quality dimensions as innate [7], but new dimensions could be added by the learning process, and learning new concepts is sometimes connected with expanding the conceptual space with new dimensions. These spatial dimensions may also be culturally dependent. The question of how 'natural' these dimensions are is avoided, but Gärdenfors is certain that they are useful from an instrumentalist point of view. He also proposes that certain neural network or statistical methods, e.g., Multi-Dimensional Scaling and Self-Organizing Maps could be used as a basis for a domain in a conceptual space [7]. The Self-Organizing Map reduces the dimensionality of the data in a systematic and meaningful way, which can be seen as moving from sub-conceptual to conceptual level.

Gärdenfors sees categories as convex regions in a conceptual space. The concepts are learned by learning a limited number of examples and by generalizing from them. The similarity of two objects can be defined as a distance between their representation points in the conceptual space.

According to him, when adopting this view of concepts, the prototype effects can be explained in the conceptual spaces. The prototypes would simply be those instances of the category that are located in the central parts of these regions. This central point would represent a possible object with the most typical features of the category, but the existence of such an object within the members of the category would not be needed.

One essential asset of the conceptual spaces theory is that it incorporates the concept of distance. This distance measure can then be used, e.g., for categorization: The perceived item belongs to the category of which the prototype is nearest to the representation of the item in conceptual spaces. The use of this kind of approach is explained in more detail in Section 2.2.2. In general, the conceptual spaces theory proposes a mediating level between sensory and symbolic levels. It provides a medium to get from the continuous space of sensory information to a higher conceptual level, where regions in it could then be associated to discrete symbols.

2.2 Formation of concepts

How does one then aquire a concept? In the framework of this Master's Thesis it is assumed that concepts are indeed learned in interaction with the world and that they are not innate [38]. Laurence and Margolis [21] suggest that one acquires a concept by assembling its features, which are often considered to correspond to sensory properties. In the following, work related to the themes of this thesis that are used for studying concept acquisition and possible mechanisms for concept learning is presented.

2.2.1 Concept learning using SOMs

Schyns [32] demonstrated how simple concepts could be learned with a modular neural network model. The model has two modules, one for categorizing the input in an unsupervised manner and another module for learning the names in a supervised mode.

The input for the Self-Organizing Map, which was used as a categorization module, was pictorial image data varied in such a fashion that there were 'prototypes', which were never directly shown to the SOM [32]. Instead, the map was fed distortions around these prototypes. In a sense, this image data could then correspond to certain 'sensory data'. The result of this experiment was that the map learned to represent the prototypes which were never fed to the system. Schyns sees that the categorization module fills the definitions of the prototype theory.

As statistical approaches, [30] and [13] can be mentioned. In these studies, an attempt was made to acquire the semantics of words from textual data. In [30], the research was conducted using generated data (the researchers generated three-word sentences themselves) and later in [13] with real word data using Grimm's tales in English as a source. In the experiments, Self-Organizing Maps and contextual information for words (preceding and following word of the target word) were used. The experiments show that based on the contextual information the target words were indeed organized in a SOM in a way that seems meaningful to us — nouns in one group, verbs in another.Words with similar usage (e.g., verbs with past tense, nouns describing humans) could also be found in smaller subgroups.

It seems, though, that a single Self-Organizing Map cannot be used for representation of the totality of concepts, but rather, various SOMs are needed for different domains. Additionally, a system to produce the per-concept feature selection for more complex concepts would be needed. See more detailed discussion on this in [19].

2.2.2 Discrimination games

A possible model of how agents could learn concepts from the environment has been proposed by Luc Steels [38] by means of discrimination games. It is based on a hypothesis that origins of meanings are based on construction and selection processes embedded in the discrimination tasks. He writes:

> Meaning is a conceptualization or categorization of reality which is relevant from the viewpoint of the agent. Meanings can be expressed through language, although they need not be. Meaning takes many forms depending on the context and nature of the situation concerned. Some meanings are perceptually grounded. Others grounded in social

interactions and others in the behavioral interaction between the agent and the environment.

The principle of perceptual grounding is that low level sensory information is used. The pressure for categorization comes from an agent's need to distinguish a target object from other objects in the context. Agents construct new segmentations of the continuous sensory space. Each object has a set of features that are values of a sensory channels.

Steels' discrimination game

In Steels' model [38], the agent plays the discrimination game alone. Initially, there are no innate features but the system knows which property of the object utilizes which sensory channel. The context, in which the game is played, includes the objects that are currently in the field of attention of the agent. One object is then selected randomly as a target or *topic* of the game. Then the feature sets for the topic and the other objects are derived. The game consists of an attempt to find possible discriminating feature sets that could separate the topic from the other objects in the context.

In Steels' model, an agent playing the games produces a discrimination tree by using emerging feature detectors for finer and finer distinctions. Each feature detector has an attribute name, a set of possible values, a function and a sensory channel. As the discrimination process advances, a hierarchical structure is formed.

If a discrimination game is unsuccessful, it implies that there are not enough distinctions to distinguish the topic from the other objects in the context. Two things are used to correct the situation: (1) If some sensory channel does not yet have feature detectors, a new one may be constructed. This is the preferred option. (2) Otherwise, an existing feature may be refined by creating a new feature detector that further segments the region covered by that feature.

If the game is successful and there is more than one possible feature set that can be used to describe the topic, a feature set is chosen based on three criteria. (1) The smallest set is preferred. Thus the least number of features are used. (2) In case of equal size, the set in which the features imply the smallest number of segmentations is chosen. Thus, most abstract features are chosen. (3) In case of equal depth of segmentation, the set of which the features have been used the most is chosen, which should encourage the development of a minimal set of features. The outcome is a hierarchical tree of segmentation of each feature space. Features that are not used at all are eliminated.

The results of their experiments show that the system is able to produce a discrimination between objects both when the number of objects is constant during the simulation and when the number of objects increases steadily [38].

Vogt's discrimination game

Vogt [48] has refined the method introduced by Steels [38] in such a way that it includes the notion of Gärdenfors' Conceptual Spaces described in Section 2.1.3. The discrimination game is used as a part of the simulation in which agents try to come up with a common vocabulary by playing language games. These language games are explained in more detail in Chapter 3.

In the simulation, each agent constructs a private ontology. Initially, the ontologies are empty. The task in the discrimination game is the same as in Steels' games: Find a category (one or more) for the object that distinguishes the topic from others in a given context, which contains some limited number (usually five) of objects. If a game fails, the agent's ontology is expanded to improve discrimination in future games.

The main difference between Steels' and Vogt's work is how categories are represented. While Steels used discrimination trees, in Vogt's work the emerging categories were represented as prototypes. These prototypes are points in an n-dimensional conceptual space [7]. The categories are defined in such a way that all the points that are nearest to the prototype presenting the category belong to that category. This is equivalent to a Voronoi tessellation of the space.

The agents use the quality dimensions of the conceptual space to construct their ontology. Vogt uses four quality dimensions, on which the agents can also perceive information: R, G, B, and S corresponding to colors red, green, blue and the shape of the object, respectively. With these quality dimensions the agents can construct different conceptual spaces, a holistic space containing all dimensions, 'RGBS', a color space 'RGB' and a shape space 'S' or a 'redgreen' space and a 'blueshape' space. Overlapping spaces are not allowed. Vogt's viewpoint is that the compositional conceptual space might be a starting point for an evolving compositional language as well.

The categorization proceeds in a similar manner as in Steels' discrimination game. An agent categorizes all objects in the context by combining features from each dimension to form a multi-dimensional category. The game succeeds if the category is distinctive: It is not a category for any other object in the visual field of the agent. If the game fails, the features of the topics are added as new exemplars of categorical features in the agent's ontology, unless they exist there already.

Again, the idea of how fine the partition of the space is depends on the objects. The finer the distinctions needed to be perceived are, the finer must be the partition of the conceptual space. When an object is categorized, it is put to the category corresponding to the nearest neighboring prototype to that object.

Chapter 3

Language emergence

In the previous chapter, the concepts and computational models for the study of their formation were discussed. In this chapter, the focus is on language and the forms of its emergence. First, the language origins are discussed briefly.

3.1 Language origins

The most controversial question regarding the origins of language is whether the language ability is based on overall cognitive abilities, or whether there is a specific language device in the brain. The question of the origins of language is not exactly at the focus of this thesis, but as the debate is a heated one, both views are summarized briefly in this section.

One still common hypothesis is that there is some kind of biological language faculty or language organ in the brain, which is separable from the general cognitive abilities [11], [28], [29]. The language learning is then only a matter of setting parameters and in this way refining innate knowledge (Universal Grammar hypothesis) [5].

An opposite view to the origins of language is that language is a culturally emergent phenomenon: thus the similarities in languages can be explained by human cognitive and social universals [42] instead of specific 'language genes'. In [39], individual adaptation, cultural evolution and self-organization are given as the basic mechanisms for language evolution. The language is seen as an autonomous evolving adaptive system maintained by a group of distributed agents without central control and it is viewed from a functional perspective and it is preserved in individuals' memories instead of in genes, and transmitted in a cultural fashion based on learning by imitation.

On the question of how the information is preserved and what shapes the language, various selectionist criteria are offered in [39]. These criteria include attempts to maximize communicative success, minimize cognitive processing and memory load. He sees that the coherence in language use, i.e., how consistently a certain utterance is used to express a certain meaning, arises in the model from self-organization, in a certain kind of positive feedback loop: If a word is successfully used, its use will be preferred in a similar situation.

3.1.1 Computational studies on language origins

Computational approaches to the study of evolved communication are numerous and only a few of them are mentioned here. These studies are very interesting examples as in them the use or need of language in the environment was taken into account.

In the study of Werner and Dyer [50] the task of the agents in the simulated environment was to find males in a situation where females were able to emit sounds and see the males, whereas the males were blind. Thus, evolving communication, the ability to hear and interpret the signals was favored in the environment.

In the study of MacLennan [23] in the field of synthetic ethology, language arises as a side effect of cooperation. Another study, by Cangelosi and Parisi [3], studies the emergence of communication in the context of how a word relates to behavior or real word phenomenon, and how language improves the fitness of the agent population. In all of these studies the need of communication arises from the criterion that the 'speaker' must be able to perceive something useful that the 'hearer' cannot. The agents are able to learn word-meaning pairs successfully, but in these studies meaning space and word/signal space are very small and pre-defined. Thus, it is rather a form of animal communication, than language in the normal sense.

For a more thorough reviews on the computational approaches to study language origins, see [4], [16] and [26]. See also [24] on different levels of language emergence.

3.2 Acquisition of word meanings

Let us now turn our attention to the question of language learning. How do language learners acquire the meaning of novel words? The difference between language acquisition and language emergence is that children living in an existing language community learn an existing language: They do not have to come up with one from the scratch[1].

One argument for the nativist view (see Section 3.1), as it is sometimes called, is the poverty of stimulus argument [8]. According to it, children cannot learn a language based on positive evidence only. Thus, language needs to be innate. For counter-arguments to the poverty of stimulus argument, see e.g., [53].

But what kind of information does a language learner receive then? An infant must receive some kind of cues which help to distinguish what the adult is talking

[1]Although, it seems that if they do not have a language, they will develop one, of which the famous Nicaraguan sign language case [34], where deaf children who weren't taught but some home-signs develop a fully functional language in two generations is a classical example.

about. These cues come both in linguistic and extra-linguistic form. Vogt lists a variety of social cues [49] available for a language learner. One, and probably the most studied one, is the joint attention. To establish the joint attention the speaker may indicate the topic of conversation by pointing or in some other manner. Thus, joint attention is an equivalent of having associative feedback: symbol and referent are available at the same time.

Another form of feedback is corrective feedback. It could be of form: "Give me a bunny." "No, not the car, the bunny." What is more important, the feedback does not need to be explicit, the child may be capable of inferring it from the context. It is also possible that neither of these cues are needed, if agents can observe the words in a context of various items. Thus, when agent perceives a word, let's say, 'fadiga' in the context of red, green and blue objects and later in the context of red, yellow and black object, it may infer that it means the red color [2].

3.2.1 Language games

The notion of language game was originally introduced by Wittgenstein [51]. In his opinion, what defines language is how it is used. To him every occasion of language use is a language game.

In a language game there is a dialogue between two agents, a speaker and a hearer, within a particular contextual setting. The language game models discussed here were introduced by Steels [37] to study how a coherent lexicon may emerge by means of cultural interactions, individual adaptation and self-organization. Within this framework, the cultural evolution of language may be studied: The evolutionary process is not at the level of subsequent agent generations but rather in subsequent language games.

[2]Of course, this requires that the previous instance was held in the memory of the agent available for later analysis.

There are currently three types of language games used within this framework. These include the observational, guessing and selfish games [49] that model different aspects of communication and language acquisition. The games have been simulated, e.g., in [6], [35] [37], [46, 47, 48] or implemented in a population of physical robots, e.g., in [40] and [44].

Observational game

In an observational game, joint attention is established between the speaker and the hearer. The learning is thought to be associative: the topic of the game (the referent in the world) and the word are presented at the same time.

In Vogt's observational games ([48], [49]), two agents are first selected randomly from the population of agents. One is assigned the role of speaker, the other is the hearer. The division of tasks is arbitrary. The speaker selects randomly one meaning from the context (in [48]) or from the shared, predefined ontology of meanings (in [49]) and informs the hearer, what the topic is by means of extra-linguistic information. In this way, the joint attention is formed.

The speaker then searches its lexicon for words that are associated with the topic of the game. Each association has a certain association score. These associations scores are adapted based on the result of the game. The association having the highest association score is selected.

If the speaker does not find any word associated with the meaning, it invents a new word and adds the word-meaning association to the lexicon with a low initial association score. Then, the selected word is uttered.

The hearer searches its own lexicon for an association in which the word matches the received word and the meaning corresponds to the topic of the game. If the hearer succeeds in finding a proper association, the game succeeds. Otherwise, it fails.

If the game is a success, both agents increase the association score of the used association and they *laterally inhibit* all competing associations. An association is competing, when either the meaning corresponds to the topic but the communicated word is different or when the word is same but the meaning is not.

If the game is a failure, the hearer adopts the word and adds the word meaning pair to its lexicon, with a low initial association score. The speaker lowers the used association. See [49] and [48] for further details.

Guessing game

A guessing game [44] is played with a finite (small) number of objects. Typically in reported experiments the size of the context or number of objects within the agent's visual field has been about five, e.g., [48]. The speaker chooses the topic of the game from this context, but does not inform the hearer about it. The speaker provides only the utterance it uses to denote the topic. The hearer must then guess, which of the objects the speaker means, hence the name of the game. The speaker produces for the hearer some corrective feedback on whether the guess was right or not. The guessing game is thus similar to reinforcement learning [41]. Both agents adapt their own lexicon according to the results of the game.

Selfish game

The third game, introduced by Vogt [44] and Smith [35], is called the selfish game. The agents have no way of knowing whether their communication was successful, as no feedback is given. Thus, the learner must infer the meanings of words from their co-occurrences in different contexts or situations. The game is called 'selfish' as in some way the speaker does not care whether the message was correctly understood.

Chapter 4

Language game approach

4.1 Introduction

The previous chapters contain more general discussion on concepts and different aspects of language and they can be seen as laying the groundwork for the more experimental work presented in following chapters. This chapter presents the approach used in this Master's Thesis and in the next chapter the experiments are presented in more detail.

There are two main problems that we are trying to address within this Thesis: The first one is how to model the conceptual learning and the second is the question of how language is acquired in such a way that a common vocabulary emerges in a population of learners (or agents). In the model presented later in more detail the learning task is divided into two phases: First, the agents develop an organized representation of the world based on the data. Next, the agents engage in series of observational games of which the basic description is given in Section 3.2.1.

In this work, the conceptual spaces approach presented in Section 2.1.3 is adopted for the modeling conceptual representations: Specifically, Self-Organizing Maps are used as models for agents' conceptual maps. (Cf. also [12] and [14] on

Self-Organizing Map as a semantic memory model.) The language acquisition process is modeled in a simulation environment. In this environment, a population of simulated agents will engage in series of language games similar to those described in section 3.2.1.

4.2 Conceptual maps

In this Thesis, the Self-Organizing Map is used as an implementation for a conceptual map. In the following, the principles of the Self-Organizing Map are presented in more detail.

4.2.1 Self-Organizing Map

The Self-Organizing Map (SOM) [18] (also called the Kohonen map or the Self-Organizing Feature Map) is a neural network model developed originally by Teuvo Kohonen in the early 1980s [17]. It produces a topological ordering by mapping the input space to an array of nodes. The purpose of the SOM is usually visualization of data sets.

Each node of the SOM consists of a prototype vector of the same dimension as the input vectors. The nodes are organized in the form of a sheet, cylinder or a toroid. Typically, the topological neighborhood is either hexagonal or rectangular as illustrated in Fig. 4.1.

The Self-Organizing Map functions according to the competitive learning principle. When an input vector is fed into the system, a prototype vector that best matches the input vector is selected. The best-matching unit (BMU) is the node with the smallest distance to the input vector in some metric. Usually the Euclidean distance is used, but other distance metrics can be used as well. If the input vector contains only partial information (some parts of the input vector are not known),

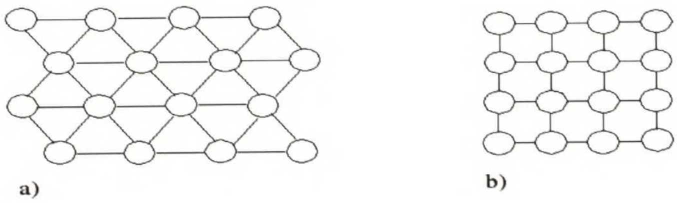


Figure 4.1: Neighborhood topologies of the Self-Organizing Map: a) hexagonal and b) rectangular.

the BMU is searched only by the existing part.

In the adaptation process, the best matching unit and its neighbors in the topological ordering are moved toward that input in the space. The degree of adaptation depends on the learning function:

$$m_i(t+1) = m_i(t) + h_{ci}(t)[x(t) - m_i(t)], \tag{4.1}$$

in which the $h_{ci}(t)$ is the neighborhood function defining how large the neighborhood is, m_i is the ith map unit, $x(t)$ the input vector, and t is the discrete time coordinate. Most commonly used is the Gaussian neighborhood function:

$$h_{ci} = \alpha(t) \cdot \exp\left(\frac{||r_c - r_i||^2}{2\sigma^2(t)}\right), \tag{4.2}$$

where $0 < \alpha(t) < 1$ is the learning-rate factor monotonically decreasing in the course of the learning and $\sigma^2(t)$ corresponds to the neighborhood radius, also decreasing monotonically in the course of the learning. The r_c and r_i are the vectorial locations on the grid. Another commonly used neighborhood function is a bubble. It is constant in the whole neighborhood (defined by neighborhood size) and zero elsewhere.

The neighborhood size and α are relatively large in the beginning of learning in order to get a rough, global ordering in the map. As the learning process continues and the values decrease, a more local ordering is achieved. Typically α is close to (but smaller than) 1 and the neighborhood size may be in the beginning of learning, half the size of the diameter of the map for the Gaussian function [18].

There are several ways to initialize the prototype vector values of the Self-Organizing Map. The prototype vectors can be initialized randomly, the initial values of the map can be taken from the available input samples, or the initialization may be linear.

There are two different training algorithms for the SOM. In sequential training the prototype vectors are adapted after each input to the SOM, as described earlier. A faster alternative is the batch training [18]. In the batch training the whole data set is presented to the map before the map prototype vectors are adapted at all. In each training step the data is then partitioned according to the Voronoi regions of the map weight vectors. Next, the new prototype vectors are calculated as

$$m_i(t+1) = \frac{\sum_{j=1}^{n} h_{ic}(t) x_j}{\sum_{j=1}^{n} h_{ic}(t)}, \tag{4.3}$$

where $c = \arg min_k \{\| x_j - m_k \|\}$ is the index of the BMU of the data sample x_j. The new prototype vector is a weighed average of the data samples, and the weight for each data sample is then neigborhood function value $h_{ic}(t)$ at its BMU c.

4.2.2 The agents' ontologies

In this work, it is thought that a domain in a Conceptual Space (see Section 2.1.3) could be represented as a Self-Organizing Map trained with observation data. In this experiment, the color data, the RGB values of color pictures, is used for training of the map. Following Gärdenfors' vocabulary, there are three *quality dimensions* in this *domain* of the Conceptual Space: the R(ed), G(reen) and B(lue). The Self-Organizing Maps are trained with the color data prior to the simulation, in which the language acquisition is studied. After the initial training of the SOM, the map is not changed. This corresponds to a situation in which a child initializes its feature representations based on natural visual data. When an object (color vector) is perceived during the simulation, it is mapped to the trained SOM by finding a unit whose Euclidean distance to the perceived input is the smallest. This map node is the best-matching unit (BMU) [18].

The objects the agents see belong to eight different categories: What the agents perceive are slightly different instances of these categories. This differs considerably from previous approaches, e.g., [49], where the meanings are presented simply as integers. Also, in [48], only distinct prototypical colors are used: There is no variation around the prototype.

The meaning of a word is taken to be a node or a group of (neighboring) nodes in the Self-Organizing Map. Thus, the word is not directly associated with 'something in the world', the referent, which in our case is the perceived data vector but to a representation: The representation of the data vector is the BMU in the map. Cf. also Vogt's discussion [45] on Peirce's semiotic triangle [25].

The association between a word and a concept is implemented by assigning a word to a certain node in the conceptual map. The mapping between words and conceptual map nodes is many-to-many. A node may have several words associated with it and a word may be associated with several nodes. We are hypothesizing that a general agreement on which word to use for which meaning emerges during the simulation.

As described earlier in this chapter, the ordering of the prototype vectors is topological in the Self-Organizing Map: Similar prototypes tend to be close to each other and those further apart are more different. It can then be assumed that if we think a concept is as an area, instances located nearby each other in a map are quite similar, and they can be labeled with the same word. Here the notion of distance, which is an essential feature of the theory of Conceptual Spaces is used. The size of the map defines how fine-grained distinctions of the data can be made.

The level of similarity can be defined with the radius R. It describes the size of the neighborhood the word-meaning associations are searched from. This is also the group of nodes that may be described with the same word and they can be seen to belong to the same category. Figure 4.2 illustrates the neighborhoods in the map. Black node is the BMU and different neighborhood sizes are marked with different colors.

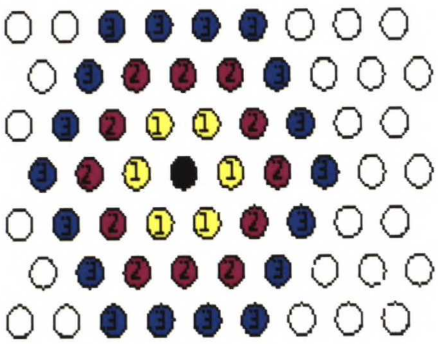


Figure 4.2: Neighborhoods of different size around the BMU (black), 1-neighborhood (yellow), 2-neighborhood (magenta), 3-neighborhood (blue).

4.3 Utterances

Each word is a discrete symbol in the simulation. A word is a string of characters generated from a simple artificial language. Following Wittgenstein's view on the purpose of language, a word is uttered when it is needed. If there does not exist a word that could be used to denote the topic of the conversation, a novel word is generated from the language.

In the experiments, a very limited artificial language is used. In this language, there are words of length four and of six characters. The alphabet contains vowels

$V = (a, e, i)$ and consonants $C = (b, c, d, f, g, h)$. In total, the alphabet consists of nine letters. Each word of this language begins with a consonant which is then followed by a vowel. The pattern is repeated either once or twice, so all the words are either of the form 'CVCV' or 'CVCVCV'. A similar approach is used in Vogt's experiments [47].

In many previous simulations, e.g., [3], the set of words that could be used was small and fixed. In these simulations the set of words is finite but open: new words can enter to the simulation, whereas the number of topics of the language games is fixed.

4.4 Observational game

As described in Section 3.2.1, in an observational game both agents know in advance what the topic of the game is. In Vogt's and Steels' robotic experiments this was accomplished by pointing, and later in simulations by using other extra-linguistic information. Our solution is that the agents are able to perceive only one 'object' at the time and this is the topic of the language game. These objects and their properties are presented in more detail in Section 5.1.3.

4.4.1 Structure of the agent

Each agent has a conceptual map based on a Self-Organizing Map and a lexicon. The lexicon contains all words that are in the agent's vocabulary, and information on which nodes of the SOM they are associated to. It also contains a counter value for the word-node pair describing how successfully a word has been used to express a meaning previously. The minimum value of the counter is zero and as the maximum value we have used is twenty.

4.4.2 The algorithm of the observational game

Each language game in the simulation proceeds in a following way.

1. Two agents are chosen randomly from the population of agents. One is assigned the role of the speaker, the other the hearer. The roles are divided arbitrarily.

2. The topic of the language game is chosen randomly from the set of topics and shown to both agents.

3. Both the speaker and the hearer search for a node in their own conceptual map that best matches the topic.

4. The speaker searches for the word that could match the topic. The search is performed in a neighborhood of the BMU defined by R, which is an integer, $R \geq 1$. The process of the word search is described later in more detail. If no possible word is found, a new word is invented and associated to the BMU. This word is communicated to the hearer.

5. The hearer searches for a set of possible words that could denote the topic. The search is performed in a similar way as in the case of speaker, but instead of one best word, all the words that are found are returned. If the word the speaker has uttered belongs to this set, the language game is considered a success, otherwise the game fails.

6. In case of a successful game, both the speaker and the hearer increase their counter for the word by one. If the uttered word was not among the labels of the BMU, it is then added to it. The maximum value of the counter is set to 20.

7. If the game fails, the speaker decreases the counter of the uttered word by one. The minimum value allowed for the counter is zero. If the speaker's

BMU node did not contain any label but the word was instead found from the neighborhood, the word is not added to the BMU node of the speaker. The hearer labels its BMU with the spoken utterance in any case.

4.4.3 The implemented word search process using the SOM

When an agent is shown the topic (the vector containing the features of the topic), it finds a prototype vector from its conceptual map (SOM) that best matches the given input vector. This prototype vector is called the best-matching unit (BMU). The prototype vectors can be assigned labels, the words. The search algorithm searches for words used to label the BMU and other nodes within the radius R neighborhood of the BMU. Even though there is not any word associated to that particular node, a *word hypothesis* can be made, if a word has been used for a similar enough object. In other words, the same word may be used to denote a group of prototype vectors.

In this way a set of possible words is assumed. From this set, the uttered word is selected: It is the one used most successfully earlier in the neighborhood of the BMU, defined by the counter values associated to each word-node pair. In case of multiple words with the same count one is selected randomly. If the set of words is empty, a new word is generated and uttered.

The hearer searches for the possible words in a similar way to the point of finding a set of words. The only difference is that there is no need to select the best word but the set of possible words is compared to the word uttered by the speaker. In general, the competing (word or node is the same) word-meaning pairs are considered to be either synonymous or polysemous, both being features of natural languages as well.

4.4.4 Cleaning of the lexicon

In the end of the simulation agents' lexicons are cleaned up by removing words that have not been used successfully during the simulation. This is done by removing all words associated to a certain node if the value of the counter is zero, but only if there is a better word hypothesis (counter value > 0) for the node in the neighborhood of that node.

4.5 Differences to previous studies

The approach presented here owes considerably to previous work, especially by Vogt [44],[49], [48]. The current model differs from the previous works in some aspects. This section summarizes briefly the main differences.

The notion of conceptual spaces is also used in [48], but the conceptual representations have been implemented in a different way. More specifically, the features of the perceived object are taken as such and used as prototypical representations of the concept. Whereas in this work, the representation of the object is the *best-matching unit* in the Self-Organizing Map.

In [48], the initial ontologies of the agents are empty, and an agent acquires the concepts during the simulation, and the acquisition of the shared vocabulary is simultaneous to conceptual learning. In the approach presented here, the learning process is divided in two phases: In the first phase the conceptual maps are trained with color data and the trained map then serves as a basis for categorization. The language emerges in the second phase, but the SOM is not changed during the simulation.

A related difference is on what the agents are able to perceive. In this work, it can be thought that the agents are able to perceive slightly different instances of eight types of objects. In both [49] and [48] the number of different objects the agents are able to perceive is considerably larger, but the notion on which instances should be considered belonging to the same category is not considered at all: All the objects should belong to different categories ultimately.

In [48], the compositionality of language is mainly studied using the Iterated Learning Model (ILM) by Kirby [15]. In the Iterated Learning Model the notion of adult and children agents is used within the context of subsequent generations: The adults are mainly the teachers and the children are the learners of the emerging language. Neither the ILM or the compositionality of language is considered in this Thesis. Also compared to [49] and [48] at this point of the study only the observational game is implemented. The guessing and selfish game are left to future research.

In Vogt's studies, e.g., [44], [49] and [48], if the language game is successful, the competing word-meaning associations (either the word or the meaning of a word-meaning pair is the same as in the winning association) are inhibited and the score of the winning association is increased. In our case only the counter of the winning association is increased and no inhibition takes place. Additionally, in our study the hearer agent produces all words that are within the radius R from the BMU whereas in Vogt's studies only the one with the biggest association score is returned. Our solution corresponds to a more natural language situation in which there are synonymous words corresponding to a meaning. The preferred words may vary for each agent: while the word used by the speaker is not the one the hearer prefers, it can still be perfectly understood.

Chapter 5

Language game experiments

This chapter reports the experiments made with the observational game described in Chapter 4. The purpose of the experimental work is two-fold: The first goal of the experiments was to verify the hypothesis that the agents are able to develop an emergent and shared lexicon by engaging in the language games, while using the SOM conceptual map model. Secondly, we are studying the association processes between the map nodes and the utterances created by agents and how the areas that are named with the same word are formed. To test how the varying parameters affect the overall learning results, experiments were conducted with different population sizes, different sizes of conceptual maps and varying the search radius, R, although the purpose of these experiments was not to find the best parameters.

5.1 Methods

The simulation program used in the experiments was constructed using Matlab[1]. The SOMs used as conceptual maps were implemented using the SOMToolbox for Matlab [43]. The program code used in the experiments is available from the author.

[1] `http://www.mathworks.com`

In all experiments, 10 simulations were run with different random seeds for 5000 language games. There were three measures used to evaluate the outcome of the simulations. The communication success was calculated after every language game. The coherence and the specificity were calculated after every 250 games, and the size of the lexicon was calculated in the end of the simulation. These measures are explained later in more detail.

In the experiments, the training of the conceptual maps was separated from language learning. A similar but considerably simpler approach was used in [49]: In that study, the meanings were represented simply as integers. Simultaneous conceptual learning and language learning was studied in [44], [46] and [48].

5.1.1 Color data

The agents' conceptual maps were trained with color data vectors. Components of the vector were R(ed), G(reen) and B(lue) values of a pixel in a color picture. The color data consisted of 10 pictures — one for each agent. The size of these pictures was 100×100 pixels. Thus for each agent, the training set contained 10000 samples.

The color pictures were created by drawing color filled ellipses and rectangles to a white background. The colors used and their equivalent RGB-values are given in Table 5.1. To get less 'spiky' distributions for each color, uniformly distributed noise was added independently to each of the three color channels (RGB) of the picture. The level of noise was set to 20% of the total color range. An example picture is shown in Figure 5.1.

5.1.2 Training of the conceptual map

In the experiments, a hexagonal map topology was used. Three different map sizes were used in the simulation experiments: small map was of size 8×6 nodes or map

Color	R(ed)	G(reen)	B(lue)
black	0	0	0
blue	0	0	1
green	0	1	0
cyan	0	1	1
red	1	0	0
magenta	1	0	1
yellow	1	1	0
white	1	1	1

Table 5.1: Colors and their equivalent RGB values.

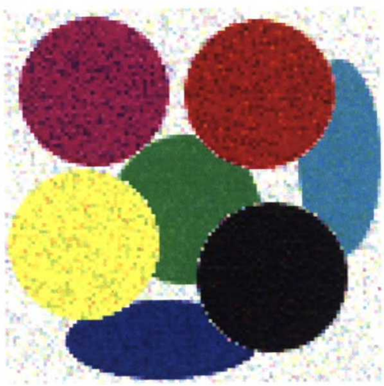


Figure 5.1: An example of a color picture used to train an agent's conceptual maps.

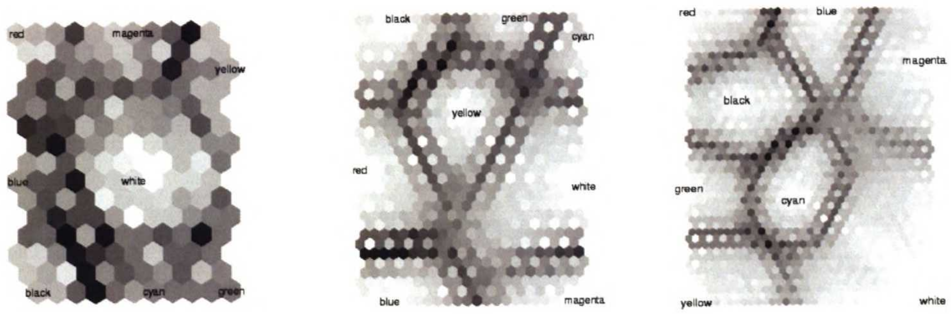


Figure 5.2: Self-Organizing maps of different size trained with the same data: left: 8×6 nodes, middle: 16×12 nodes, and right: 24×18 nodes (random initialization).

units, the medium size map was 16×12 nodes and the large map was 24×18 nodes. The maps were initialized randomly.

The maps were trained in a batch training mode, explained in section 4.2.1. The length of the training varied with the size of the map. The length of the training was the default 'long' training, defined in the SOMToolbox functions [43] and it was divided into two phases. The length of the rough training in epochs was

$$t_{rough} = 16n/l, \tag{5.1}$$

and for fine tuning the used length was

$$t_{fine} = 64n/l, \tag{5.2}$$

where n is the number of the nodes in the map and l is the length of the training data, which is 10000 in this case. The training length is always at least 1 epoch. The initial radius for the training was $rad_{ini} = \max(1, \max(mapsize)/2)$ and the final radius for training is always $rad_{fin} = 1$. The value of the learning factor α (see Eq. 4.2 on page 25) was kept at 0.5 for the rough training and 0.05 for the fine training phase. The parameters for each map size are listed in Table 5.2.

Map type	map size	t_{rough} in epochs	t_{fine} in epochs	rad_{ini}	rad_{fin}
Small map	[8 6]	1	1	4	1
Medium map	[16 12]	1	1.23	8	1
Large map	[24 18]	1	2.76	12	1

Table 5.2: Parameters for training the SOMs.

Figure 5.2 shows the U-matrices of the three SOMs trained with the same data and illustrates the differences between them. In these U-matrices, the light colored nodes mean that prototypes associated with the nodes are near each other in the input space. The darker the color, the further apart the prototypes are in the input space. A clear partitioning to eight different regions can be seen. One can also notice that the larger the map, the clearer the separation to different regions in the map is. The color names were added to the maps later to illustrate the organization of the map: The BMUs were searched for each color vector, presented in Table 5.1, and the corresponding color name was added as a label to that node. This kind of labeling was not used during the simulations.

5.1.3 Language game topics

A set of objects that the agents are able to perceive was created for the topics of language games. This was achieved by creating an additional picture in a similar way as the pictures used for training the agents' conceptual maps. For the purpose of limiting the computational workload, the size of this picture was limited to 20×20 pixels. Thus, there were 400 different topics for the games. For each game, the topic was chosen randomly from this group.

5.1.4 Evaluation measures

When the main purpose of the work is to find out *whether the agents are able to communicate successfully*, one should properly define what is meant by successful communication. The reader should bear in mind that within this framework the use of communication is more limited than in human populations. Thus, the definition of successful communication is also more limited. In the previous Chapter (p. 30), the definition for the successful language game was given: If the word used by the speaker to denote the topic is found from the word set containing all possible words the hearer associates with the topic, the game is considered as a success.

The overall communication success (CS), the pure outcome of the language games, was measured in a similar way as in [6], [44] and [49]. It was calculated as an average number of correctly played games in the past 100 games or less, if no 100 games had been played yet. It was calculated after every language game.

If we now pursue further to the realms of language, there are other issues that contribute to the usefulness of the language, than only whether the language game was successful or not. One of them is the language coherency, i.e., a certain word is coherently used to denote a certain meaning in the agent community. Thus, a high coherence level indicates that the agents have developed a shared lexicon that they are using.

Coherence is a population measure. It is the rate in which agents would produce a certain word to express a meaning. The coherence measure calculation used in this work is taken from [6]. For each topic, a fraction of agents that has the same word as a preferred word is calculated and the maximum fraction is taken. This is then averaged over all topics. If an agent does not have a word to express a meaning, the coherence is set to zero. The calculation of the coherence is illustrated in Table 5.3.

The initial experiments were conducted using only the two aforementioned

	ρ_1	ρ_2	ρ_3	ρ_4	ρ_5
a_1	cihahe	fadiba	ceda	dibi	dibagi
a_2	cihahe	fadiba	ceda	hadafe	cada
a_3	fadiba	fadiba	ceda	hadafe	dibagi
a_4	dibagi	fadiba	dibi	dibi	dibagi
max. freq	2	4	3	2	3
coherence	0.5	1.0	0.75	0.5	0.75

Table 5.3: Calculation of coherence for referents ρ_1–ρ_5 and for agents $a_1 - a_4$.

measures. When examining those earlier results, it was noticed that if agents use only one word to describe every possible referent, the communication success and the coherence are still high. If we only evaluate the evolving language based on communication success and coherence, a language consisting of only one word would be considered useful, although it lacks the quality of being able to differentiate between the referents. Thus, an additional measure, specificity, was introduced.

In this work, a specificity measure developed by De Jong [6] is used. In [6], De Jong describes: "[S]pecificity indicates to what degree the words an agent uses determine the referent that is the subject of communication". The specificity decreases if two meanings are referred to with the same word. In other words, specificity describes the amount of polysemy in the lexicon: the higher the specificity, the less polysemy there is. In this thesis, the specificity based on preferred words [6] is used. De Jong introduces also a measure of specificity based on entropy, but for the purpose of this thesis the straight-forward use of the preferred-words based specificity is sufficient.

For each agent, A_i,, the specificity, $spec(A_i)$, is calculated from the following formula:

$$spec(A_i) = \frac{n_s^2 - \sum_{k=1}^{n_s} f_k}{n_s^2 - n_s}, \tag{5.3}$$

	ρ_1	ρ_2	ρ_3	ρ_4	ρ_5	$\sum f_k$	Specificity
a_1	1	1	1	1	1	5	1.0
a_2	1	1	1	1	1	5	1.0
a_3	2	2	1	1	1	7	0.9
a_4	2	1	2	2	2	9	0.8

Table 5.4: Calculation of specificity for agents according to [6] for referents ρ_1–ρ_5 and for agents $a_1 - a_4$.

where n_s is the number of referents, and f_k is frequency of the word related to the concept, which describes how many referents the word is associated to. The specificity of the population, $spec$, is then defined as the average specificity of the agents:

$$spec = \frac{\sum_{i=1}^{n_a} spec(A_i)}{n_a}, \tag{5.4}$$

where $spec(A_i)$ is the specificity of an agent and n_a is the number of agents. Additionally, if there is not a word to denote a certain referent (or topic), it means that the referent cannot be separated from other referents. Now, it can be thought that the referent is associated with all the other referents, and it has the frequency $f_k = n_s - 1$, where n_s is the number of referents. The calculation of the specificity for the words of Table 5.3 is illustrated in Table 5.4.

Compared to the work by De Jong [6], where mapping between perceptions and 'meanings' is one-to one, in this work it is many-to-one: The agents are performing some kind of rough categorization as well. At this point, the specificity measure was introduced to show differences between different choices of parameters, differences that otherwise could not have been perceived at all. Thus, it was sufficient to calculate specificity only for the eight prototypical color vectors presented in Table 5.1.

In the end of each simulation run, the average size of lexicon was also calculated. In the lexicons, there were also words having a zero counter value, which means that they were not used successfully at all in the course of the simulation, i.e., an agent had come up with a word, but it had not been used successfully and another word had been preferred since. To better show the difference in lexicon sizes, the average size of the lexicon was calculated both before and after the removal of the non-used words. The average lexicon size was calculated as a mean of individual agent lexicon sizes.

5.2 Results

This section summarizes the results of the experiments. Altogether, the simulations were run using small, medium and large maps. For the small map, the radii $R = 1$ and $R = 2$ were used, and for medium sized map radii $R = 1$, $R = 2$, and $R = 4$ were used. The calculation using the large maps turned out to be very time-consuming. Thus, simulation runs were conducted only with the radius $R = 2$. For all experiments, the results are averaged over 10 simulation runs. All simulations were run for the population sizes of 2, 4, 6, 8 and 10 agents except for the large map, where only population sizes of 2, 4 and 6 agents were used, due to the computational load. In the following, the significant results are highlighted. Figures showing the results from all experiments are given in Appendix A.

5.2.1 First results with varying population size

First, overall results with varying population size are presented. In these experiments, the middle-sized map was used and the search radius was set to $R = 2$.

The results are presented in Figure 5.3. The communication success (a) climbs quickly close to the maximum value of 1.0. The communication success level 1.0 indicates that each of the previous 100 language games ended successfully. The larger the population is, the longer it takes to reach the maximum level, but even in the case of 10 agents the level of 0.9 is reached after approximately 1250 games.

The coherence level (b) increases also quite rapidly. In case of population size of 10 agents, the coherence level reaches 0.86, whereas with smaller population sizes it is 0.9 or higher. Thus, as simulation advances, the agents begin to use more and more the same word to denote the same referent, thus forming a shared vocabulary. As pointed out earlier, the coherence does not say anything whether the agents are using the same word to denote *each* referent. The specificity (c), rises over 0.9 already after 250 games with each population size. Thus, there seems to be no polysemy in this case: The agents are using a separate word for each prototypical color (see Table 5.1).

The average size of the lexicon (d) stays between 10 and 17, rising only a little as a function of the population size. The whiskers describe the amount of standard deviation in the average values. The size of the lexicon, all words included, rises highly as a function of the population size.

These first results seem promising. They clearly show that the agents can develop a shared lexicon to denote the objects they perceive. The size of the population seems only to affect on how quickly the communication success and coherence levels grow. The reason behind this is that in larger populations, it takes longer for the whole population to develop a common vocabulary, as in each language game, there are always only two agents playing.

5.2.2 Varying the search radius

Next, the results from simulations with different search radii are presented. Now we are comparing results where the population size and the size of the map were kept constant. The middle-sized map with 16×12 nodes is used, and there are six

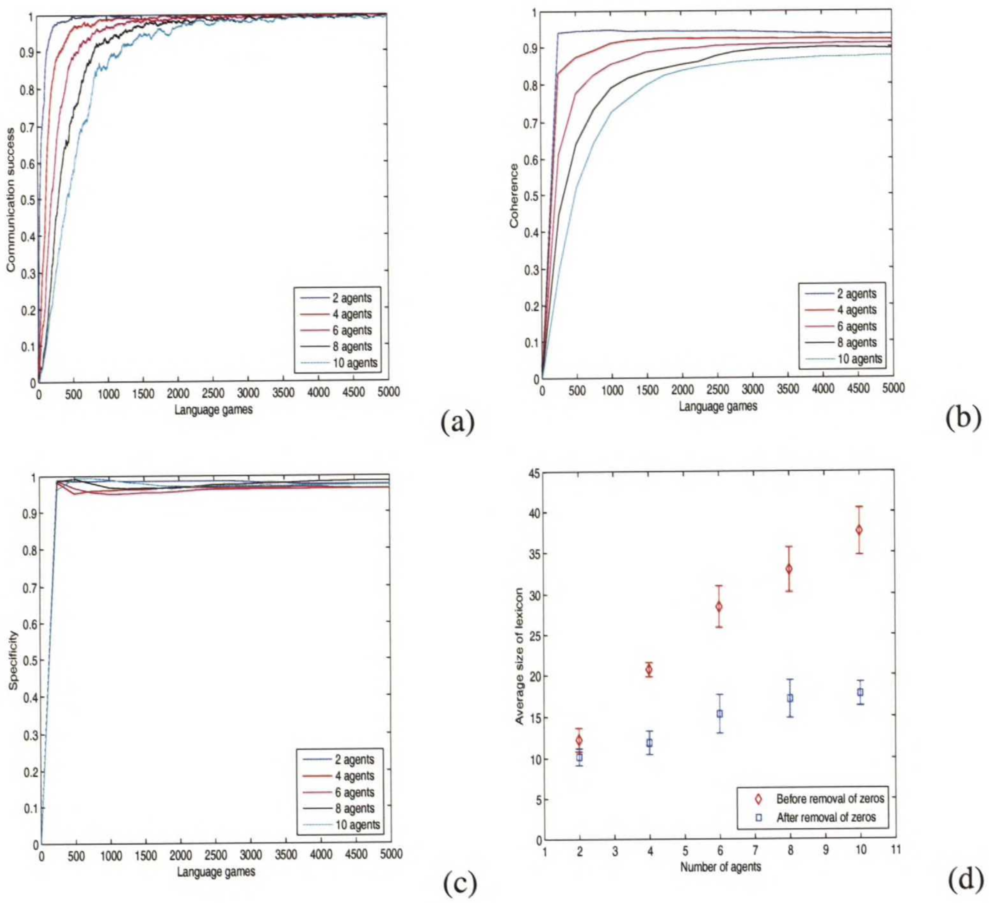


Figure 5.3: Communication success (a), coherence (b), specificity (c) and the lexicon size (d) for varying population size, when $R = 2$ and the map size is 16×12.

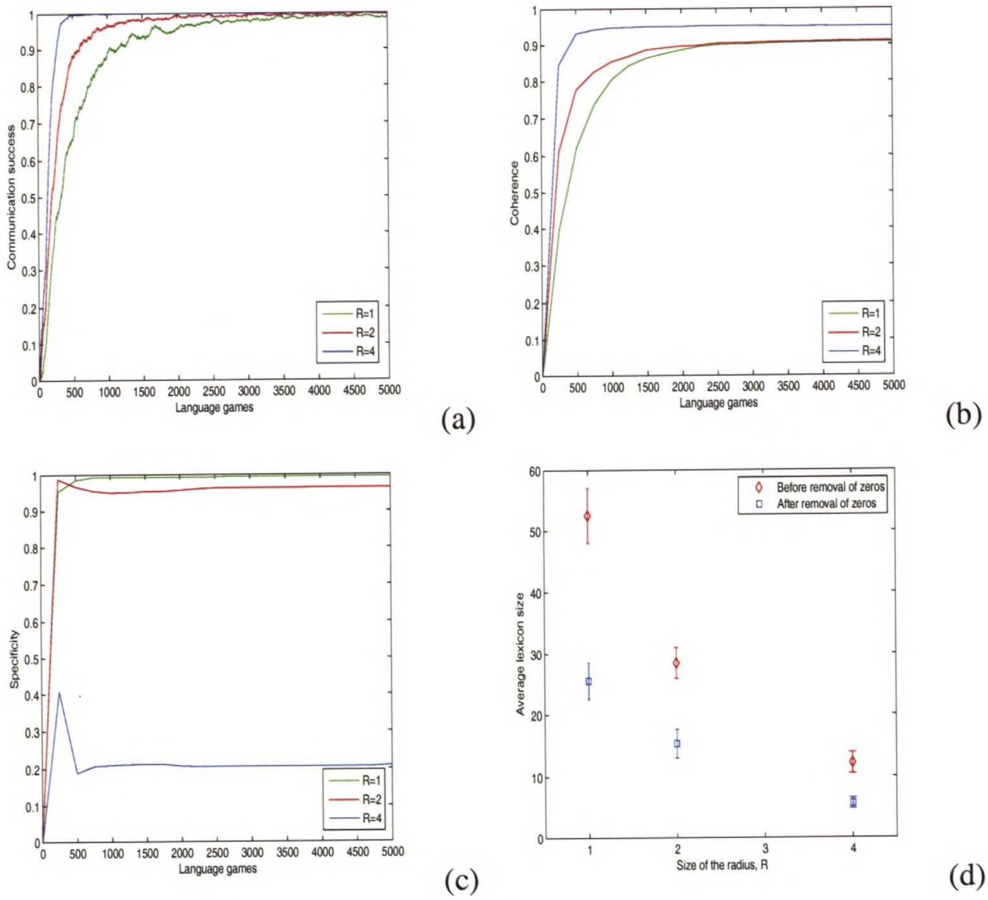


Figure 5.4: Communication success (a), coherence (b), specificity (c) and average lexicon size (d) with varying the search radius when middle-sized (16×12) map was used and the population size was six agents.

agents in the population. The results are presented in Figure 5.4.

Let us first inspect the communication success (a). With all search radii the communication success reaches the maximum of 1.0. It seems that the larger the radius, the faster the communication success level rises. This is due to the fact that when the search space is small, the possibility to create a new word for a referent increases. This in turn makes the vocabulary larger and as a consequence, it takes longer for the whole population to learn all the words and to use them successfully.

The coherence (b) seems also to climb the fastest with the largest search radius, and the final level is 0.95 whereas it is 0.9 when using $R = 1$ or $R = 2$: Again, the agents seem to be using coherently the same word to name the same referent.

Now the specificity measure (c) reveals something interesting. With the smaller radii, the specificity is again over 0.95, and climbs there after only 500 language games. But when using the largest radius, $R = 4$, the specificity first rises to the level of 0.4 and then drops to the level of 0.3 as the simulation advances: This indicates that in the beginning of the simulation there is some variation in the names for the language game topics. In the course of simulation few words are gaining more and more popularity. And if we now look at the average lexicon size (d), it seems to be very small (approximately 5). Remember, the lexicon size only tells how many words have been successfully used at some point during the simulation: Not every one of those is used in the end of the simulation.

Example conceptual maps are shown in Figure 5.5. They present U-matrices labeled with the words used during the simulation. The one on the left is taken from one of the ten simulation runs with the population size of six agents, where $R = 1$ and the figure in the right presents a conceptual map from a simulation run again with six agents, where $R = 4$. The visual interpretation of these figures may be hard since the language is made up by the agents themselves. Thus, they are shown for general interest only. It seems that in the map on the right there is one dominant word, 'bihi', used to label almost everything. On the left-hand side there are different words that seem to be dominant in different areas: 'fiha' in the bottom left corner, 'bihe', in the bottom middle, 'ficeca' in the bottom right corner, 'fifeba' in the middle left, 'gaci' in the middle right, 'gedede' in the top right, 'figifi' in the top middle. In the top left corner there are various words. By only inspecting the map it is impossible to say, which one of these is used the most.

The low specificity value does not affect the communication success at all. The agents just seem to have one common word to denote everything. To explain

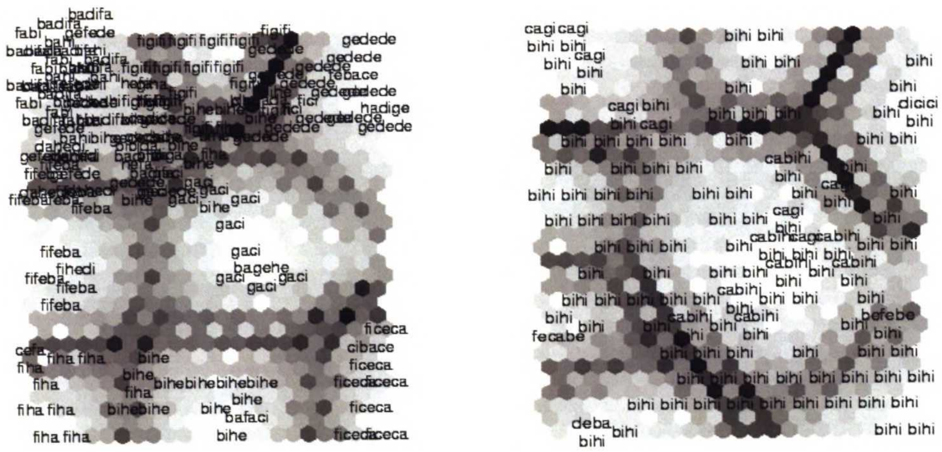


Figure 5.5: Two example conceptual maps of size 16×12 from simulations of 6 agents. On left: $R = 1$, on right: $R = 4$.

this behavior, we will have to look again at the naming process used in the simulation. As explained in Section 4.4.3, the agent is searching for a name associated to the node that best matches the perceived language game topic. This node is called the best matching unit (BMU). The word search process is extended to all the nodes that are within the R-neighborhood (see Fig. 4.2) of the BMU node. Conversely, if there is one node that is labeled (the word is used in a language game) with a certain word, all the nodes within the R-neighborhood can be also labeled with that word, if they happen to be BMU for another topic in some following language game. Then again, all the nodes in the R-neighborhood can become labeled with the same word.

In the experiments presented here, the Self-Organizing Map size was 16×12. The hexagonal lattice was used and the search radius was $R = 4$. We can think that there is a 'first' word which becomes associated to one node in the map. If the radius is $R = 1$, there would be maximum 6 other nodes that could be associated with the same word in the next game (the immediate neighbors of the node, on the side or corner of the map there are of course less neighbors). If the radius is $R = 2$, there are at the most 18 nodes that can be associated with the word. And when the radius is $R = 4$, there maximum number of nodes that can be associated with the word is 60. As the total number of nodes in the map is 192, the nodes that can be associated with the same word cover almost a third of the whole map. Thus, when another node further away from the first BMU becomes associated with the same word and this happens multiple times, all the map nodes can easily become associated with only one word quite fast.

5.2.3 Different mapsizes

Next we will compare the results of simulations with different map sizes. Again the population size used was six agents. The search radius was kept at $R = 2$. The results are presented in Figure 5.6. Both the communication success (a) and the coherence (b) level rise slowest with the biggest map. Eventually the level of 1.0 is reached in case of the communication success with all map sizes. The coherence level reached is 0.9. The results seem natural: while the search radius is small in proportion to the map size, it takes longer to assign names for all the nodes in the map. Consequently, it takes longer until the common vocabulary is achieved. The specificity (c) and the average lexicon size results are as expected: When the map size is small (and the search radius is large compared to it), the specificity again drops in the same way as when we were using the middle sized map and $R = 4$. (Compare to the results with small map and $R = 1$ in Fig. A.1, where specificity level reaches 1.0 very quickly.)

As mentioned earlier, the calculation was very heavy when the large map was used and the processing took very long. Our results do not reveal any specific advantages of using the large map for the data we have been using. But are there any differences in performance between the small and the middle-sized map, if the search radii are selected appropriately according to the map size? Figure 5.7

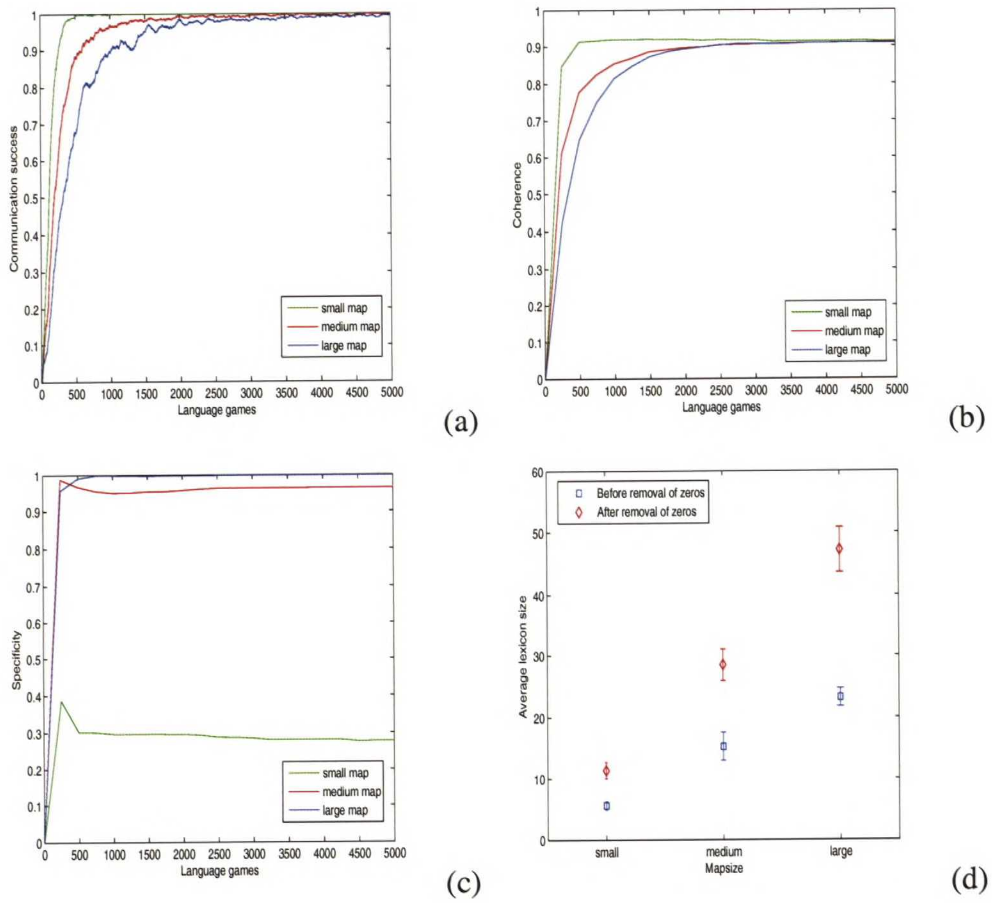


Figure 5.6: Communication success (a), coherence (b), specificity (c) and average lexicon size (d) when comparing three different map sizes, using population size of 6 agents and the search radius $R = 2$.

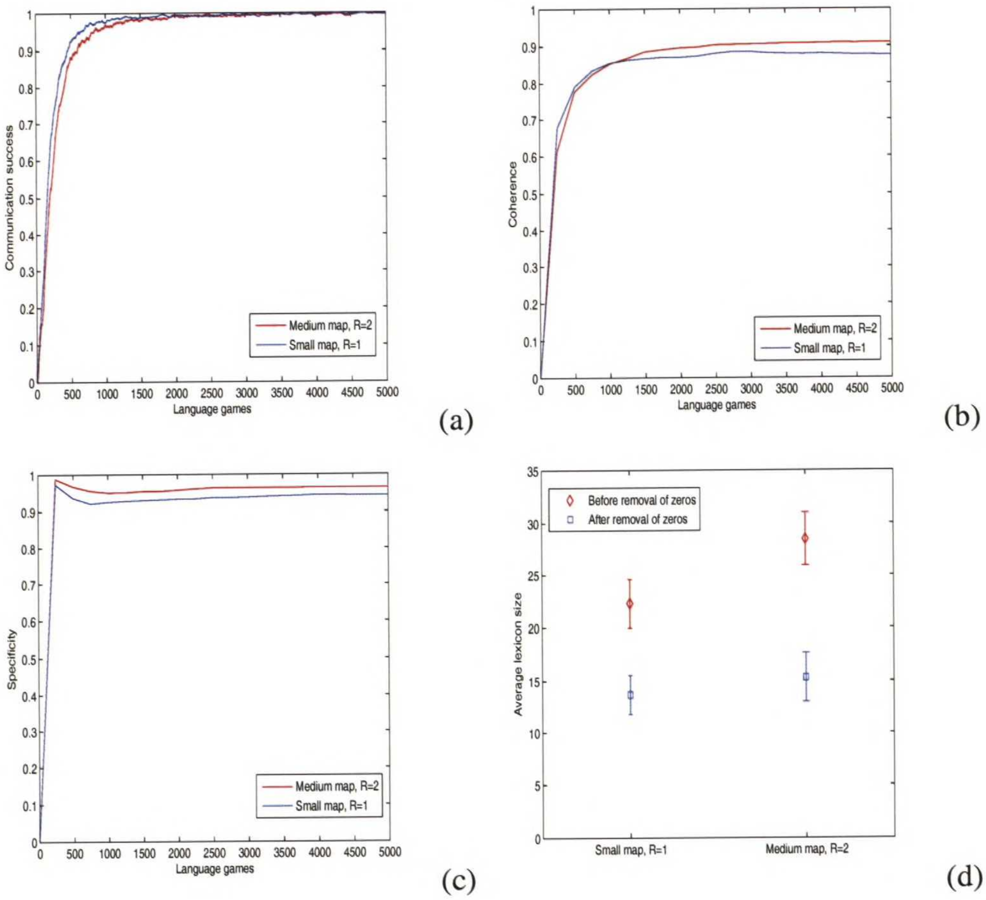


Figure 5.7: Comparison between a small map with $R = 1$ and middle-sized map with $R = 2$: Communication success (a), coherence (b), specificity (c) and average lexicon size (d). There were six agents in the population.

presents a comparison for simulations using small maps with search radius $R = 1$ and middle-sized maps with $R = 2$. Again, the agent population consisted of six agents. Based on these results, there does not seem to be any difference in the performance: In both cases, communication success rises to the maximum level of 1.0, there is a slight difference in the coherence levels and the specificity levels are similar. The only considerable difference is between the average lexicon sizes, if all created words are taken into account. But when only the successfully used words are taken into account, the difference is again very small. Thus, based on these observations, it seems that the smaller map might be the better choice if one wants to limit the computational workload without affecting the results.

5.3 Discussion on the results

In this section, some comments on the experimental results are presented. General discussion on the domain of the thesis is presented in Chapter 6. The system seems certainly to be working: The agents are able to map the perceived topics to the maps and associate utterances to them. According our definition of successful communication, the agents are also able to communicate successfully and develop a shared lexicon based on adaptation.

In these experiments the choice of the map sizes and the search parameter R was quite arbitrary. If R is too small compared to the map size, every node is labeled with a different word and if the R is too large, the same word is used to denote all referents. Thus, one should be able to find a good middle grounds. It seems that the use of small maps and $R = 1$ or middle-sized maps and $R = 2$ suits best for the data used in the experiments.

As we have seen, the agents may begin to use only one word to denote all the referents. This seems to be due to the too large R in relation to the map size. It seems also important to point out that in the context of the observational game, the language is not necessarily needed: It is somewhat redundant as the hearer agent always knows what the speaker is referring to. In guessing and selfish games there is a crucial need to distinguish between different objects, whereas in the observational game there is no such pressure that would force the agents to develop separate words.

The averaging over 10 games was done in a similar fashion as in [49]. Further statistical analysis would have been important to see whether there were large inter-simulation variations that could not be now perceived at all. It would have been interesting to add a measure to describe the global lexicon formation. This measure could have been used to see how many words are used at each point during the simulation. Now one could only draw some conclusions based on the conceptual maps and the individual agent lexicons in the end of the simulation.

In the experiments conducted for this thesis, the natural 'borders' created between the regions (see Fig 5.2) of the SOM, i.e., the distances visualized by the U-matrix, were not used at all. Instead, the search was based on the neighborhood of the BMU only (Fig. 4.2). By using the knowledge on how near or far the neighboring node is, one might be able to create groups of words which correspond better to the perceived world structure (pre-trained maps).

Chapter 6

Discussion

6.1 Conclusions

In this Master's thesis work some aspects of language acquisition and conceptual modeling have been considered. In the field of conceptual modeling, the Conceptual Spaces Theory by Gärdenfors [7] has been adopted. The theory provides a medium between the symbolic level of words and the sensory level of 'raw' sensations. The notion of distance provides a possibility to make graded conceptual system: The more prototypical instances of a concept can be seen as more central than the less-prototypical instances of the category.

The Self-Organizing map was used as a basis for the conceptual map of an agent. Each agent's conceptual map was trained with color data prior to the learning. The observations were then mapped to the conceptual map and labeled. Within this framework, the pre-trained SOM seems a suitable basis for the conceptual map as it can be used to reduce the representational complexity of the input data.

As a model for shared vocabulary acquisition, different types of language games were discussed in this thesis. A computer simulation to model one of them, the observational game, was implemented based on the work presented in [37], [44], and [49]. The main difference to the language game simulations presented in those studies is that the conceptual maps of the agents were based on the Self-Organizing map. The acquisition of concepts was not simultaneous to language learning in the simulations, but the conceptual maps were trained with color data prior to the language game simulations. The experiments were conducted using different sizes of Self-Organizing Maps and search radii. The population size was also varied between two and ten agents.

The results of the experiments show clearly that when using the observational game model and the SOM-based conceptual maps (1) the agents learned to communicate successfully on the topics of the games and (2) a shared lexicon was developed during the simulations.

6.2 Future work

6.2.1 Other language game models

In this Master's Thesis, only the observational game was implemented. It would be interesting to implement the Guessing game and the Selfish game described in Section 3.2.1 as well to see how the proposed conceptual map implementation behaves with them. As pointed out in Section 5.3, the language is somehow redundant in the observational game framework: Both the speaker and the hearer know for sure what the topic of the game is and there is no need to have distinct words to separate different objects. It would be interesting to study if the need to be able to identify the topic from a group of objects is a pressure enough to prevent the agents of calling all referents with the same word.

Wittgenstein [51] listed various kinds of language games that are used in different situations. A further and a very interesting continuation to the research would be to create a simulation environment in which the three language game types presented within this thesis would be used when appropriate. Of course, what defines the appropriate use is again a difficult question, where pragmatics should be taken into account.

6.2.2 Simultaneous learning of concepts and words

In this Master's Thesis work, the learning from visual data and the acquisition of the vocabulary was divided in to two distinct phases: The Self-Organizing map was not changed at all during the simulated language games. Anyhow, Vogt [48] and others [20], [51], [7] argue that semantics of languages are a product of co-development of language and meaning in embodied interaction of individuals in their environment: Learning simpler concepts and associating words with them helps with the acquisition of more complex concepts. Thus, the language is seen as a means for creating a coherent conceptual structure. Within the simulation framework described here, the direct implementation is not possible, as the continuous training of the Self-Organizing Map has problems that are not yet solved.

6.2.3 Use of multiple domains of Conceptual Spaces

In Chapter 4, the use of multiple maps for different domains was discussed briefly. In this model, a more complex concept say 'apple', would have properties in different domains. E.g., 'green' (or 'yellow', or 'red') in color domain, 'round(ish)' in shape domain, 'sweet' in taste domain etc. Gärdenfors [7] argues that which of these properties would be important to the one holding the concept, would depend on the context. The context-dependency would then cause some properties to be more salient in that context. It is possible that these saliencies could be modeled with some kinds of weights. Possibly, the research could be expanded further to somewhat complex concepts: To those with properties extending to different domains of Conceptual Spaces.

6.2.4 Representation of action

So far we have discussed only the relations between objects in the world and their links to the language. If we expand the scope further, how could action representation be modeled within the conceptual spaces? How would the action be then represented in the map is a good question. Perhaps it could be modeled as some kind of sensory information in an appropriate domain or appropriate domains. Gärdenfors also addresses the representation of action very briefly in the context of conceptual spaces. He describes action as dynamic properties of objects. To him the action could be modeled using for example forces that are applied to body parts during the action, and one could then have a conceptual space for these dynamical forces. He seems sure that even if the analysis of the action is tedious, the functional properties can be, in principle, explained from the more basic properties.

Bibliography

[1] M. Aydede. The language of thought hypothesis. `http://plato.stanford.edu/archives/fall2004/entries/language-thought/`, 2004. In The Stanford Encyclopedia of Philosophy, Fall 2004 edition.

[2] L. W. Barsalou. Situated simulation in the human conceptual system. *Language and Cognitive Processes*, 18(5/6):513–562, 2003.

[3] A. Cangelosi and D. Parisi. The emergence of a 'language' in a population of evolving neural networks. *Connection Science*, 10(2):83–97, 1998.

[4] A. Cangelosi and D. Parisi, editors. *Simulating the Evolution of Language*. Springer, London, 2002.

[5] N. Chomsky. Three models for the description of language. *IEEE Transactions on Information Theory*, 2(3), 1956.

[6] E. De Jong. *Autonomous Formation of Concepts and Communication*. PhD thesis, Vrije Universiteit Brussel, June 2000.

[7] P. Gärdenfors. *Conceptual spaces: The Geometry of Thought*. MIT Press, 2000.

[8] E.M. Gold. Language identification in the limit. *Information and Control*, 10:447–474, 1967.

[9] S. Harnad. Psychophysical and cognitive aspects of categorical perception: A critical overview. In S. Harnad, editor, *Categorical Perception: The Groundwork of Cognition*. Cambridge University Press, New York, 1987.

[10] S. Harnad. The symbol grounding problem. *Physica D*, 42:335–346, 1990.

[11] M.D. Hauser, N. Chomsky, and W.T. Fitch. The faculty of language: What is it, who has it and how did it evolve? *Science*, 298, 2002.

[12] T. Honkela. Neural nets that discuss: A general model of communication based on self-organizing maps. In S. Gielen and B. Kappen, editors, *Proceedings of International Conference On Artificial Neural Networks (ICANN-93)*, pages 408–411, Amsterdam, 1993. Springer-Verlag.

[13] T. Honkela, V. Pulkki, and T. Kohonen. Contextual relations of words in Grimm tales, analyzed by self-organising map. In *Proceedings of International Conference on Artificial Neural Networks, ICANN-95*, pages 3–7, 1995.

[14] T. Honkela and J. Winter. Simulating language learning in community of agents using self-organizing maps. Publications in Computer and Information Science Report A71, Helsinki University of Technology, 2003.

[15] S. Kirby. Spontaneous evolution of linguistic structure: an iterated learning model of the emergence of regularity and irregularity. *IEEE Transactions on Evolutionary Computation*, 5(2):102–110, 2001.

[16] S. Kirby. Natural language from artificial life. *Artificial Life*, 8(2):185–215, 2002.

[17] T. Kohonen. Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43:59–69, 1982.

[18] T. Kohonen. *Self-Organizing Maps*. Springer Series in Information Sciences. Springer, 1995.

[19] K. Lagus, A. Lehtola, M. Määttä, and S. Virkki. Emergence of structure and relations in conceptual representations. In T. Honkela, K.I. Hynnä, K. Lagus, and J. Särelä, editors, *Adaptive and Statistical Approaches in Conceptual Modeling*, Publications in Computer and Information Science, Report A75. Helsinki University of Technology, 2005.

[20] G. Lakoff. *Women, Fire, and Dangerous Things*. University of Chicago Press, Chicago, US, 1987.

[21] S. Laurence and E. Margolis. Concepts and cognitive science. In E. Margolis and S. Laurence, editors, *Concepts: Core Readings*. MIT Press, Cambridge, MA, 1999.

[22] J. Locke. *An essay concerning human understanding*. Oxford University Press, 1690/1975. Available at `http://www.arts.cuhk.edu.hk/Philosophy/Locke/echu/`.

[23] B. MacLennan. Synthetic ethology: An approach to the study of communication. In C.G. Langton, C. Taylor, J. Doyne Farmer, and S. Rasmussen, editors, *Artificial Life II: The Second Workshop on the Synthesis and Simulation of Living Systems*. Addison-Wesley, 1992.

[24] B. MacWhinney. Models of the emergence of language. *Annual Review of Psychology*, 49:199–227, 1998.

[25] C.S. Peirce. *Collected Papers*, volume 1. textlog.de, 1931. Available online at `http://www.textlog.de/charles_s_peirce.html`.

[26] A. Perfors. Simulated evolution of language: a review of the field. *Journal of Artificial Societies and Social Simulation*, 5(2), 2002.

[27] R. Pfeifer and C. Scheier. *Understanding Intelligence.* MIT Press, 1999.

[28] S. Pinker. Language as an adaptation to the cognitive niche. In M. Christiansen and S. Kirby, editors, *Language evolution: States of the Art*. Oxford University Press, New York, 2003.

[29] S. Pinker and R. Jackendoff. The faculty of language: What's special about it. *Cognition*, 95(2):201–236, 2005.

[30] H. Ritter and T. Kohonen. Self-organizing semantic maps. *Biological Cybernetics*, 61:241–254, 1989.

[31] E. Rosch. Principles of categorization. In A. Collins and E. E. Smith, editors, *Readings in Cognitive Science, A Perspective from Psychology and Artificial Intelligence*, pages 312–322. Morgan Kaufman Publishers, San Mateo, California, 1988.

[32] P. Schyns. A modular neural network model of concept acquisition. *Cognitive Science*, 15:461–508, 1991.

[33] J. R. Searle. Minds, brains, and programs. *The Behavioral and Brain Sciences*, 3, 1980.

[34] A. Senghas, M. Coppola, E.L. Newport, and T. Supalla. Argument structure in Nicaraguan sign language: The emergence of grammatical devices. In E. Hughes, M. Hughes, and A. Greenhill, editors, *Proceedings of the Boston University Conference on Language Development*, volume 21, pages 550–561, Boston, 1997. Cascadilla Press.

[35] A. D. M. Smith. Establishing communication systems without explicit meaning transmission. In J. Kelemen and P. Sosik, editors, *Proceedings of the 6th European Conference on Artificial Life, ECAL 2001*, Berlin, Heidelberg, 2001. Springer-Verlag.

[36] L. Steels. The artificial life roots of artificial intelligence. *Artificial life*, 1(1), 1994.

[37] L. Steels. Emergent adaptive lexicons. In P. Maes, editor, *SAB96*, Cambridge, MA, 1996. MIT Press.

[38] L. Steels. Perceptually grounded meaning creation. In M. Tokoro, editor, *ICMAS96*. AAAI Press, 1996.

[39] L. Steels. The synthetic modeling of language origins. *Evolution of Communication*, 1(1):1–34, 1997.

[40] L. Steels and P. Vogt. Grounding adaptive language games in robotic agents. In C. Husbands and I. Harvey, editors, *Proceedings of the Fourth European conference on Artificial Life*, Camridge, MA and London, 1997. MIT Press.

[41] R.S. Sutton and A.G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA, 1998.

[42] M. Tomasello. Language is not an instinct. *Cognitive Development*, 1995.

[43] J. Vesanto, J. Himberg, E. Alhoniemi, and J. Parhankangas. *The SOM Toolbox for Matlab 5*. Helsinki University of Technology, Neural Networks Research Center, P.O.Box 5400, 02015 HUT, April 2000.

[44] P. Vogt. *Lexicon Grounding in Mobile Robots*. PhD thesis, Vrije Universiteit Brussel, November 2000.

[45] P. Vogt. Anchoring of semiotic symbols. *Robotics and Autonomous Systems*, 2003.

[46] P. Vogt. Grounded lexicon formation without explicit reference transfer: who's talking to who? In W. Banzaf, T. Christaller, P. Dittrich, J.T. Kim,

and J. Ziegler, editors, *Advances in Artificial Life - Proceedings of the 7th European Conference on Artificial Life (ECAL)*. Springer-Verlag, 2003.

[47] P. Vogt. THSim v3.2: The Talking Heads Simulation tool. In W. Banzaf, T. Christaller, P. Dittrich, J.T. Kim, and J. Ziegler, editors, *Advances in Artificial Life - Proceedings of the 7th European Conference on Artificial Life (ECAL)*. Springer-Verlag, 2003.

[48] P. Vogt. The emergence of compositional structures in perceptually grounded language games. *Artificial Intelligence*, in press.

[49] P. Vogt and H. Coumans. Investigating social interaction strategies for bootstrapping lexicon development. *Journal of Artificial Societies and Social Simulation*, 6(1), 2003.

[50] G. Werner and M. Dyer. Evolution of communication in artificial organisms. In *Artificial Life II*, volume C. Langton and C. Taylor and D. Farmer and S. Rasmussen, pages 659–687. Addison-Wesley Pub, CA, 1992.

[51] L. Wittgenstein. *Filosofisia tutkimuksia*. WSOY, 1953/1999. Finnish translation of Philosophische Untersuchungen.

[52] T. Ziemke. On the role of robot simulations in embodied cognitive science. *AISB Journal*, 1(4), 2003.

[53] W. Zuidema. How the poverty of the stimulus solves the poverty of the stimulus. In Suzanna Becker, Sebastian Thrun, and Klaus Obermayer, editors, *Advances in Neural Information Processing Systems 15 (Proceedings of NIPS'02)*, Cambridge, MA, 2003. MIT Press.

Appendix A

Results of all experiments

A.1 Small map

A.2 Middle-sized map

A.3 Large map

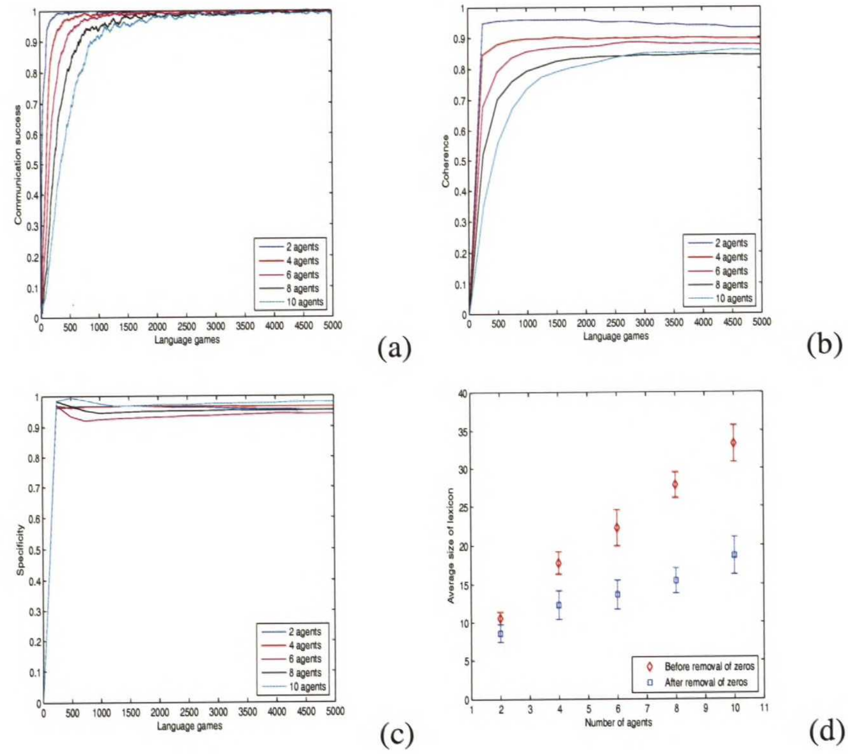


Figure A.1: Communication success (a), coherence (b), specificity (c) and the average lexicon size (d) for varying number of agents, when R = 1 and when used maps were small.

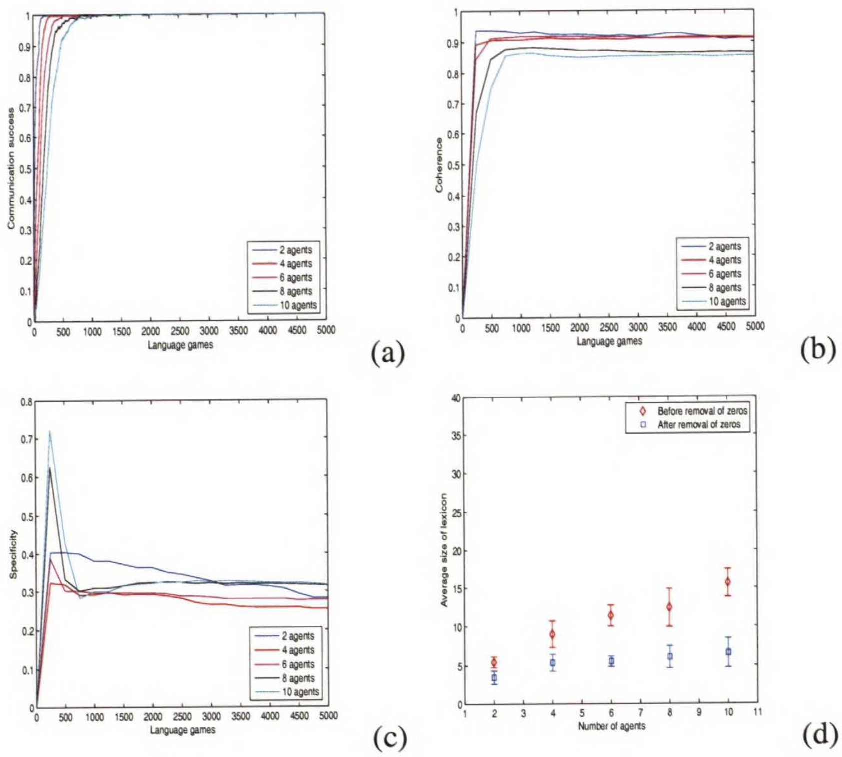
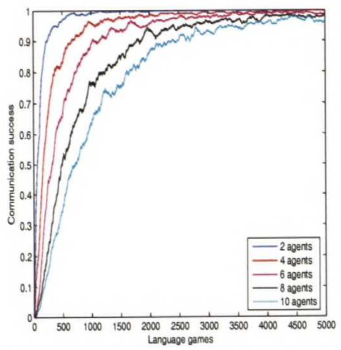
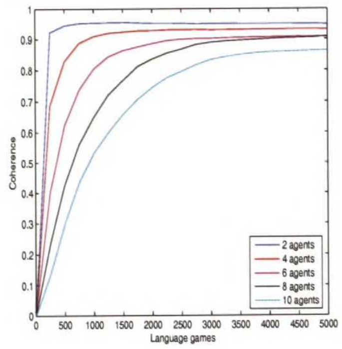


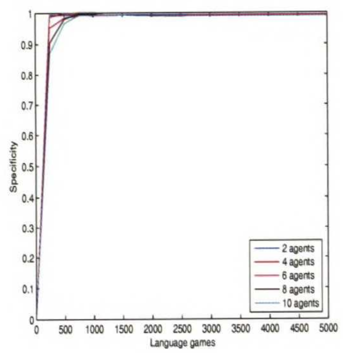
Figure A.2: Communication success (a), coherence (b), specificity (c) and the average lexicon size (d) for varying number of agents, when R = 2 and when the used conceptual maps were small.



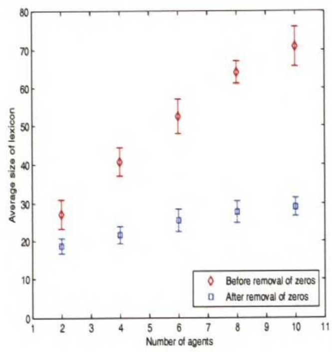
(a)



(b)



(c)



(d)

Figure A.3: Communication success (a), coherence (b), specificity (c) and the average lexicon size (d) for varying number of agents, when $R = 1$ and map size 16×12.

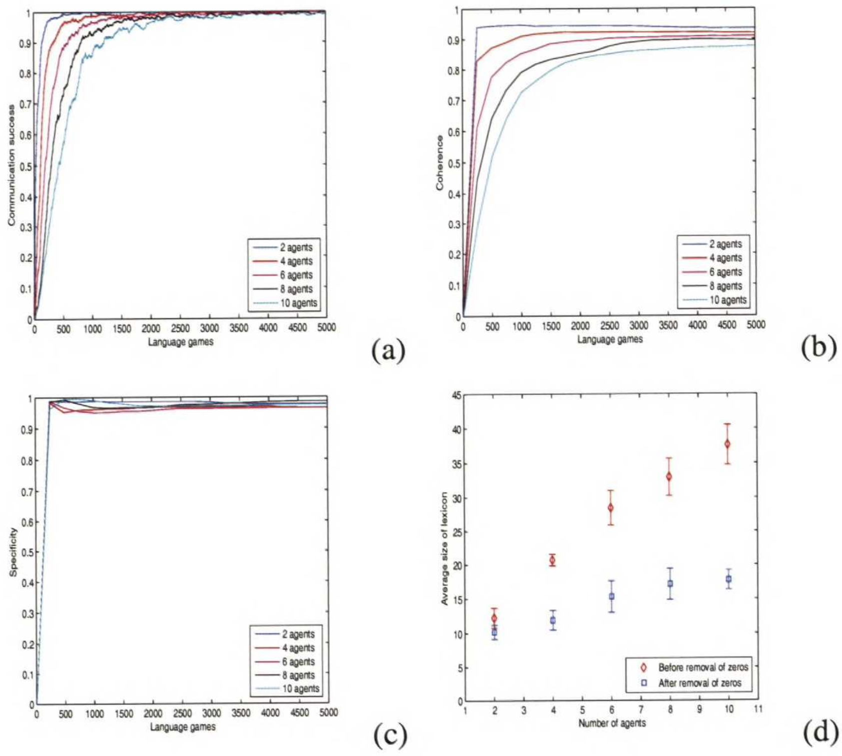


Figure A.4: Communication success (a), coherence (b), specificity (c) and the lexicon size (d) for varying population size, when $R = 2$ and the map size 16×12.

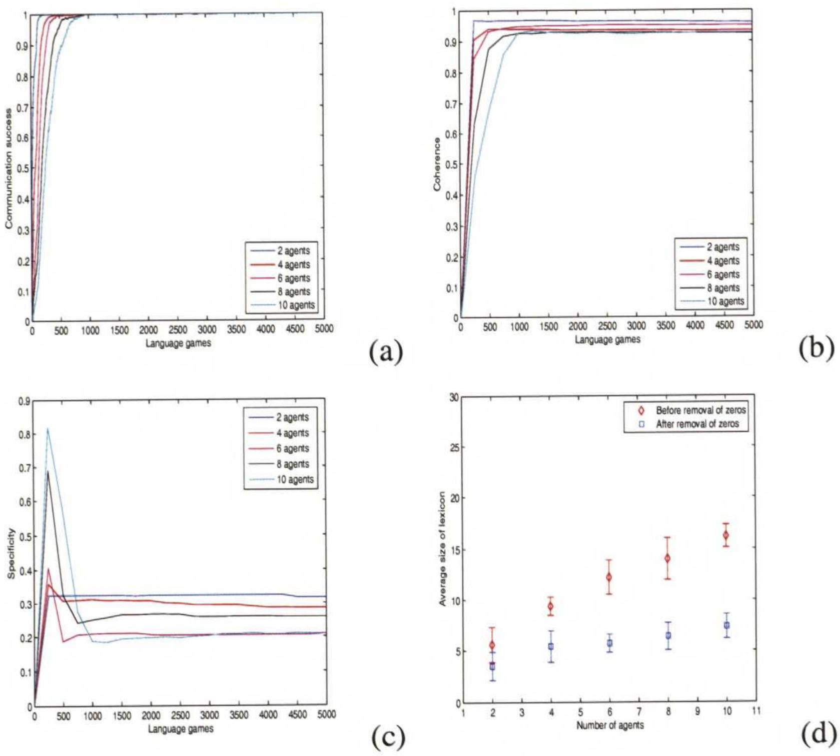
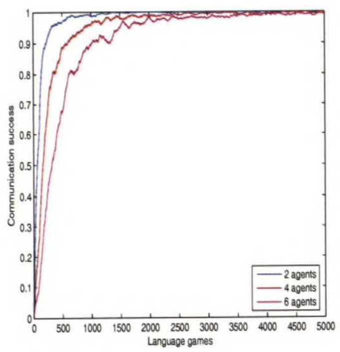
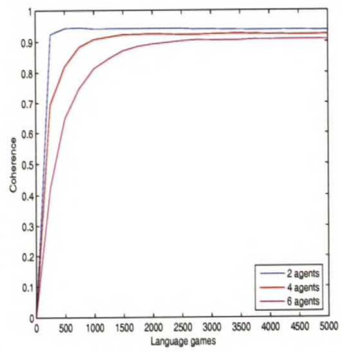


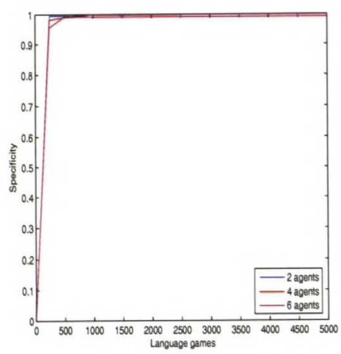
Figure A.5: Communication success (a), coherence (b), specificity (c) and the lexicon size (d) for varying population size, when $R = 4$ and map size 16×12.



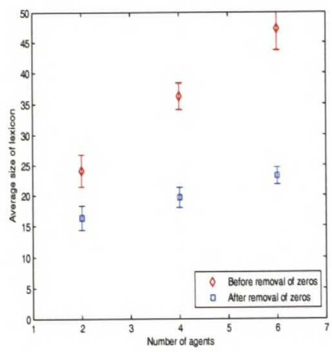
(a)



(b)



(c)



(d)

Figure A.6: Communication success (a), coherence (b), specificity (c) and the average lexicon size (d) for varying number of agents, when R = 2 and when the used conceptual maps were large.